RIESZ POTENTIALS AND NONLINEAR PARABOLIC EQUATIONS

TUOMO KUUSI AND GIUSEPPE MINGIONE

To Neil Trudinger for his 70th birthday

Abstract. The spatial gradient of solutions to nonlinear degenerate parabolic equations can be pointwise estimated by the caloric Riesz potential of the right hand side datum, exactly as in the case of the heat equation. Heat kernels type estimates persist in the nonlinear case.

Contents

1. Results

In this paper we are going to consider nonlinear, possibly degenerate parabolic equations whose model is given by

$$u_t - \operatorname{div}\,(|Du|^{p-2}Du) = \mu \tag{1.1}$$

where μ denotes in the most general case a Borel measure with finite total mass. Although the kind of problems considered are nonlinear our goal is to provide a suitable *linear potential theory* aimed at describing, in a sharp way, the regularity properties of the gradient Du in terms of those of μ. More precisely, our description shows that sharp gradient pointwise estimates can be given in terms of classical Riesz caloric potentials of the right hand side μ. We will see that, surprisingly enough, bounds similar to those that hold for the heat equation

$$u_t - \triangle u = \mu \tag{1.2}$$

actually hold for solutions to (1.1) and, more in general, for solutions to quasilinear equations of the type

$$u_t - \operatorname{div} a(x,t,Du) = \mu \qquad \text{in } \ \Omega_T = \Omega \times (-T,0)\,, \tag{1.3}$$

provided suitable, actually optimal, regularity assumptions are made on the partial map $(x,t) \mapsto a(t,x,z)$. Here $\Omega \subset \mathbb{R}^n$ is an open subset, $n \geq 2$ and $T > 0$. Specifically, we shall consider a Carathéodory vector field $a\colon \Omega_T \times \mathbb{R}^n \to \mathbb{R}^n$ which is C^1-regular in the third variable and satisfying the following parabolicity and

continuity conditions:

$$
\left\{
\begin{array}{c}
|a(x,t,z)| + |\partial_z a(x,t,z)|(|z|^2+s^2)^{1/2} \le L(|z|^2+s^2)^{(p-1)/2} \\
\nu(|z|^2+s^2)^{(p-2)/2}|\xi|^2 \le \langle \partial_z a(x,t,z)\xi,\xi\rangle \\
|a(x,t,z)-a(x_0,t,z)| \le L\omega(|x-x_0|)(|z|^2+s^2)^{(p-1)/2}
\end{array}
\right. \tag{1.4}
$$

whenever $z,\xi \in \mathbb{R}^n$, $x,x_0 \in \Omega$, $t \in (-T,0)$, where $0 < \nu \le L$ are positive numbers and $p \ge 2$. The symbol $\omega(\cdot)$ denotes a modulus of continuity meaning that $\omega\colon [0,\infty) \to [0,1]$ is a nondecreasing concave function such that $\omega(0) = 0$. In the following $s \ge 0$ is a parameter that will be used to distinguish the degenerate case ($s = 0$), which covers the model equation in (1.1), from the non-degenerate one ($s > 0$); the analysis made in the following will see no difference between these two cases. In the rest of the paper we shall assume that the partial map

$$
x \to \frac{a(x,t,z)}{(|z|^2+s^2)^{(p-1)/2}}
$$

is Dini-continuos in the sense that

$$
\int_0^1 \omega(\varrho)\,\frac{d\varrho}{\varrho} < \infty\,. \tag{1.5}
$$

This assumption is optimal for the estimates we are going to derive in the following; see the comments at the beginning of Section 1.3. We anyway remark that, everywhere in this paper, we only assume measurability of the partial map $t \to a(x,t,z)$, in other words we assume that time coefficients are merely measurable. Yet, in this paper we shall always consider the case

$$
p \ge 2
$$

as the case $p < 2$ has already been treated elsewhere [26] and involves an analysis which is different and somewhat simpler than the one which is necessary here. For more notation we refer the reader to Section 2 and to the rest of this introductory section.

Remark 1.1 (On the notion of solution)**.** Throughout the paper, when considering weak solutions to (1.3) and unless otherwise stated, we shall mean energy weak solutions. An energy weak solution u belongs to the parabolic energy space, i.e.

$$
u \in C^0(-T,0;L^2(\Omega)) \cap L^p(-T,0;W^{1,p}(\Omega))\,, \tag{1.6}
$$

and it is a distributional solution to (1.3) in the sense that

$$
-\int_{\Omega_T} u\varphi_t\,dx\,dt + \int_{\Omega_T} \langle a(x,t,Du), D\varphi\rangle\,dx\,dt = \int_{\Omega_T} \varphi\,d\mu \tag{1.7}
$$

holds whenever $\varphi \in C_c^\infty(\Omega_T)$. In view of the available approximation theory we shall assume that $\mu \in L^1$ without loss of generality, while upon letting $\mu\lfloor_{\mathbb{R}^{n+1}\setminus\Omega_T} = 0$, we shall finally consider the case

$$
\mu \in L^1(\mathbb{R}^{n+1})\,. \tag{1.8}
$$

Assumptions (1.6) and (1.8) will be then finally removed in Section 1.2. There we shall deal with solutions to general measure data problems, where both (1.6) and (1.8) are not any longer in force. In other words, we purse the usual path that consists of first deriving a priori estimates for more regular problems and solutions, and then recovering the general case by approximation. Notice that under the assumptions (1.6) and (1.8) by standard density arguments the identity in (1.7) remains valid whenever $\varphi \in W^{1,p}(\Omega_T) \cap L^\infty(\Omega_T)$ has a compact support.

1.1. **Intrinsic and explicit Riesz potential estimates.** Very recently, in [27, 28, 38] for the case $p \geq 2$ and in [13] for the subquadratic one, it has been shown that, surprisingly enough, the regularity theory of possibly degenerate quasilinear elliptic equations of the type

$$-\operatorname{div} a(Du) = \mu$$

completely reduces to that of standard Poisson equation

$$-\triangle u = \mu \tag{1.9}$$

up to the C^1-level, i.e. up to the gradient continuity. Moreover, in some sense the regularity theory *can be actually linearized via Riesz potentials.* In particular, the gradient of solutions can be pointwise bounded via classical Riesz potentials exactly as it happens for solutions to (1.9), i.e., the inequality

$$|Du(x_0)|^{p-1} \leq c\mathbf{I}_1^{|\mu|}(x_0, r) + c\left(\fint_{B(x_0,r)} (|Du| + s)\, dx\right)^{p-1} \tag{1.10}$$

holds for a.e. point x_0, where

$$\mathbf{I}_1^{|\mu|}(x_0, r) := \int_0^r \frac{|\mu|(B(x_0, \varrho))}{\varrho^{n-1}} \frac{d\varrho}{\varrho}$$

denotes the standard truncated Riesz potential of $|\mu|$. Our aim is to build a related theory for general degenerate parabolic problems of the type in (1.1) and (1.3). The main challenge here is to match the anticipated a priori Riesz potential estimate with the inhomogeneous nature of equations such as (1.1); it will be indeed part of the work to find *the proper formulation, suited to the geometry of the equations considered,* making this possible. We also remark that, even when applied to the stationary case, our results turn out to be more general than those contained in [27, 28] since the equations considered here are also allowed to have coefficients. The ultimate outcome is that, once again, the (spatial) gradient regularity theory of solutions to (1.3) can be unified in a natural way with the one of the usual heat equation (1.2). The analysis here unavoidably involves the concept of the *intrinsic geometry*, introduced and widely employed by DiBenedetto [9, 46, 2]. According to this principle, the lack of scaling (for $p \neq 2$) of equations as

$$u_t - \operatorname{div}(|Du|^{p-2}Du) = 0 \tag{1.11}$$

can be locally rebalanced by performing the regularity analysis of the solution on certain special cylinders adapted to the solution itself, indeed called *intrinsic parabolic cylinders*. More precisely, instead of using standard parabolic cylinders

$$Q_r(x_0, t_0) := B(x_0, r) \times (t_0 - r^2, t_0)\,, \tag{1.12}$$

one uses *cylinders whose time-length is stretched accordingly to the size of the gradient on the cylinder itself.* In other words, one is lead to consider cylinders of the type

$$Q_r^\lambda(x_0, t_0) := B(x_0, r) \times (t_0 - \lambda^{2-p}r^2, t_0)\,, \qquad \lambda > 0\,, \tag{1.13}$$

on which it simultaneously happens that a condition of the type

$$\fint_{Q_r^\lambda(x_0,t_0)} |Du|\, dx\, dt \lesssim \lambda \tag{1.14}$$

is satisfied. The use of the word intrinsic stems from the very basic fact that the parameter λ appears on both sides of (1.14). Ultimately, this has the effect of rebalancing the local anisotropic character of the equation allowing for proving *homogeneous regularity estimates:* in some sense, the equation (1.11) looks like the heat equation when considered on $Q_r^\lambda(x_0, t_0)$. For instance, when considering

standard parabolic cylinders, for solutions to (1.11) it is only possible to prove bounds of the type

$$\sup_{Q_{r/2}(x_0,t_0)} |Du| \leq c(n,p) \fint_{Q_r(x_0,t_0)} (|Du|+s+1)^{p-1}\, dx\, dt\,, \tag{1.15}$$

whose lack of homogeneity precisely reflects that of the equation. In this sense the previous estimate is natural. When instead considering intrinsic cylinders with (1.13)-(1.14) being in force, *estimates become dimensionally homogeneous*:

$$c(n,p)\left(\fint_{Q_r^\lambda(x_0,t_0)} (|Du|+s)^{p-1}\, dx\, dt\right)^{1/(p-1)} \leq \lambda \Longrightarrow |Du(x_0,t_0)| \leq \lambda\,. \tag{1.16}$$

Both (1.15) and (1.16) are basic results of DiBenedetto [9] while we just remark that intrinsic geometries are nowadays a basic and common tool to treat degenerate parabolic equations [1, 19, 20].

The previous considerations, together with (1.15)-(1.16), are actually the starting point for proving the desired potential estimates. Let us see how. Beside the usual caloric (truncated) Riesz potentials built upon standard parabolic cylinders as in (1.12), that is

$$\mathbf{I}^\mu_\beta(x_0,t_0;r) := \int_0^r \frac{|\mu|(Q_\varrho(x_0,t_0))}{\varrho^{N-\beta}}\,\frac{d\varrho}{\varrho}\,, \qquad 0<\beta\leq N:=n+2\,, \tag{1.17}$$

we introduce the *intrinsic Riesz potentials* as

$$\mathbf{I}^\mu_{\beta,\lambda}(x_0,t_0;r) := \int_0^r \frac{|\mu|(Q^\lambda_\varrho(x_0,t_0))}{\varrho^{N-\beta}}\,\frac{d\varrho}{\varrho}\,, \tag{1.18}$$

where N is called, as usual, the parabolic dimension. Note that in such a way we have $\mathbf{I}^\mu_\beta(x_0,t_0;r) \equiv \mathbf{I}^\mu_{\beta,1}(x_0,t_0;r)$. At this stage the word "intrinsic" merely refers to the fact that the additional parameter λ has been considered in the definition in (1.18), while at the moment no local linkage with solutions of the type in (1.14) has been considered yet. This will come in a few moments: indeed, the right way to give an intrinsic formulation of the linear potential bounds is inspired by (1.16) and it is given in the following:

Theorem 1.1 (Intrinsic Riesz potential bound)**.** *Let u be a solution to* (1.3) *under the assumptions* (1.4)-(1.5) *and* (1.8)*. There exist a constant $c>1$ and a radius $R_0>0$, both depending only on $n,p,\nu,L,\omega(\cdot)$, such that the following implication holds:*

$$\begin{aligned} c\mathbf{I}^\mu_{1,\lambda}(x_0,t_0;r) + c\left(\fint_{Q_r^\lambda(x_0,t_0)} (|Du|+s)^{p-1}\, dx\, dt\right)^{1/(p-1)} &\leq \lambda \\ \Longrightarrow |Du(x_0,t_0)| &\leq \lambda \end{aligned} \tag{1.19}$$

whenever $Q_r^\lambda(x_0,t_0)\subset\Omega_T$, (x_0,t_0) is Lebesgue point of Du, and $r\leq R_0$. When the vector field $a(\cdot)$ is independent of x, no restriction occurs on r, i.e., $R_0=\infty$.

Note that, as in a sense it was a priori required, (1.19) allows to recover (1.16) when $\mu=0$; this is a first sign of the fact that (1.19) is the "correct intrinsic extension" of (1.10). As a matter of fact Theorem 1.1 implies a gradient linear potential estimate involving standard Riesz potentials. Surprisingly enough, this is of the same type as the one which holds for the standard heat equation; moreover, when $\mu=0$, this reduces (1.15). We indeed have the following:

Theorem 1.2 (Riesz potential bound in classic form)**.** *Let u be a solution to* (1.3) *in Ω_T under the assumptions* (1.4)-(1.5) *and* (1.8)*. There exists a constant c, depending only on $n, p, \nu, L, \omega(\cdot)$, such that*

$$|Du(x_0,t_0)| \leq c\mathbf{I}_1^{|\mu|}(x_0,t_0;r) + c\fint_{Q_r(x_0,t_0)} (|Du|+s+1)^{p-1}\,dx\,dt$$

holds whenever $(x_0,t_0) \in \Omega_T$ is a Lebesgue point of Du and whenever $Q_r(x_0,t_0) \subset \Omega_T$ is a standard parabolic cylinder such that $r \leq R_0$; here R_0 is the radius introduced in Theorem 1.1.

For the case $p = 2$ the previous estimate recovers the main parabolic result in [12]. An immediate consequence of Theorem 1.2 is the following global bound via classical, non-truncated caloric Riesz potentials:

Corollary 1.1. *Let u be a weak solution to the equation*

$$u_t - \operatorname{div} a(t, Du) = \mu \qquad \text{in } \mathbb{R}^{n+1} \tag{1.20}$$

under the assumptions (1.4)*; moreover, assume that the global integrability $u \in L^{p-1}(-\infty, t_0; W^{1,p-1}(\mathbb{R}^n))$ holds for $t_0 \in \mathbb{R}$. There exists a constant c, depending only on n, p, ν, L, such that the upper bound*

$$|Du(x_0,t_0)| \leq c\int_{\{t<t_0\}} \frac{d|\mu|(x,t)}{d_{\rm par}((x,t),(x_0,t_0))^{N-1}}$$

holds whenever (x_0,t_0) is a Lebesgue point of Du.

Remark 1.2. In Theorem 1.1 $d_{\rm par}(\cdot)$ denotes the standard parabolic distance in $\mathbb{R}^{n+1}$, which is defined by

$$d_{\rm par}((x,t),(x_0,t_0)) := \max\left\{|x-x_0|, \sqrt{|t-t_0|}\right\}.$$

The previous result shows that Theorems 1.1 and 1.2 play the role of the usual representation formulae via heat kernels for solutions to the heat equation. In recent years there has been a large activity devoted to the understanding the extent to which heat kernel estimates are still valid when passing to more general settings, as for instance Lie groups and manifolds [11, 41]. In this paper we are interested, in a dual but yet related way, to see the extent to which estimates as those implied by well-behaving heat kernels can be recovered in the nonlinear degenerate setting. Our results also connects to a recent fact observed in [34], and concerning the p-superharmonicity of linear Riesz potentials.

The proof of Theorem 1.1 opens the way to an optimal continuity criterion for the gradient still involving only classical Riesz potentials and that, as such, is again independent of p.

Theorem 1.3 (Gradient continuity via linear potentials)**.** *Let u be a solution to* (1.3) *in Ω_T under the assumptions* (1.4)-(1.5) *and* (1.8)*. If*

$$\lim_{r\to 0} \mathbf{I}_1^{\mu}(x,t;r) = 0$$

locally uniformly w.r.t. (x,t), then Du is continuous in Ω_T.

An important corollary involves Lorentz spaces:

Corollary 1.2 (Lorentz spaces criterion)**.** *Let u be a solution to* (1.3) *in Ω_T under the assumptions* (1.4)-(1.5)*. If $\mu \in L(N,1)$, that is if*

$$\int_0^\infty |\{(x,t)\in\Omega_T \,:\, |\mu(x,t)|>\lambda\}|^{1/N}\,d\lambda < \infty\,,$$

then Du is continuous in Ω_T.

Corollary 1.2 substantially improves the ones in [23] claiming only the boundedness of the gradient under the assumption $\mu \in L(N,1)$; see also [6] for a related global boundedness result in the elliptic case. It might be interesting to note how the above result naturally extends to the parabolic case the classical gradient continuity results valid in the elliptic case, starting from those available for the Poisson equation $-\triangle u = \mu$ in domain of $\mathbb{R}^n$. For this it is known that the condition $\mu \in L(n,1)$ is a sufficient one for the gradient continuity. This is in turn related to, and indeed implied by, a classical result of Stein [42] that claims the continuity of a function f whenever its distributional derivatives belong to $L(n,1)$. Corollary 1.2 gives the precise nonlinear parabolic analog of such classical facts. As expected, the space dimension n is replaced by the parabolic one $N = n+2$, which is naturally associated to the standard parabolic metric.

Preliminary to the proof of the continuity criterion, there is another result which claims the VMO gradient regularity under weaker assumptions on the measure μ.

Theorem 1.4 (VMO gradient regularity)**.** *Let u be a solution to* (1.3) *under the assumptions* (1.4)-(1.5) *and* (1.8)*. If $\mathbf{I}_1^{\mu}(x,t;r)$ is locally bounded in Ω_T and if*

$$\lim_{r\to 0} \frac{|\mu|(Q_r(x,t))}{r^{N-1}} = 0 \tag{1.21}$$

locally uniformly in Ω_T w.r.t. (x,t), then Du is locally VMO in Ω_T, that is

$$\lim_{R\to 0} \sup_{r\leq R, Q_r\subset Q'} \fint_{Q_r} |Du - (Du)_{Q_r}|\, dx\, dt = 0$$

for every open subset $Q' \Subset \Omega_T$.

1.2. **General measure data problems.** Solutions to measure data problems are usually found by approximation procedures via solutions to more regular problems. These are of the type

$$(u_h)_t - \operatorname{div} a(x,t,Du_h) = \mu_h \in L^\infty\,, \qquad h \in \mathbb{N}\,, \tag{1.22}$$

where (1.6) holds for u_h and μ_h is usually a convolution of μ. The point is that solutions to measure data problems do not belong, in general, to the energy space. This section is also aimed at justifying that we may actually work under the *apparently additional assumption* (1.6). More precisely, the exact definition of SOLA, is given in the following:

Definition 1 ([4, 21, 22])**.** A SOLA (Solution Obtained as a Limit of Approximations) to (1.3) is a distributional solution $u \in L^{p-1}(-T,0;W^{1,p-1}(\Omega))$ to (1.3) in Ω_T, such that u is the limit of solutions

$$u_h \in C^0(-T,0;L^2(\Omega)) \cap L^p(-T,0;W^{1,p}(\Omega))$$

to equations as (1.22) in the sense that $u_h \to u$ in $L^{p-1}(-T,0;W^{1,p-1}(\Omega))$, $L^\infty \ni \mu_h \to \mu$ weakly in the sense of measures and such that

$$\limsup_h |\mu_h|(Q) \leq |\mu|(\lfloor Q \rfloor_{\text{par}}) \tag{1.23}$$

for every cylinder $Q = B \times (t_1,t_2) \subseteq \Omega_T$, where $B \subset \Omega$ is a bounded open subset.

In the right hand side of (1.23) appears the symbol $\lfloor Q \rfloor_{\text{par}}$, which denotes the parabolic closure of Q defined in (2.1) below. For more on this kind of solutions see Remark 5.1 below; in particular, requiring (1.23) is neither unnatural nor restrictive. Our estimates remain valid for SOLA and, in fact, the following holds:

Theorem 1.5. *The statements of Theorems 1.1, 1.2, 1.3 and 1.4 continue to hold for SOLA $u \in L^{p-1}(-T,0;W^{1,p-1}(\Omega))$ to* (1.3)*, under the only assumptions* (1.4)-(1.5)*. As a consequence, the results in Corollaries 1.1 and 1.3 hold for SOLA as well.*

1.3. **Comparison with nonlinear estimates.** Theorem 1.1 improves the previously potential estimates via nonlinear potentials [25], bringing them *to the desired optimal level*. Based on elementary dimension analysis we conjecture that the result of Theorem 1.1 cannot be improved by the use of any other nonlinear potential. Theorem 1.1 is optimal also with respect to the regularity assumed on the coefficients dependence $x \mapsto a(x,\cdot)$, that is (1.5). Indeed, already in the linear elliptic case

$$\operatorname{div}(\tilde{a}(x)Du) = 0\,,$$

Dini-continuity of elliptic coefficients matrix $\tilde{a}(x)$ is essential in order to get gradient boundedness. Merely continuous coefficients are not sufficient to ensure that the gradient belongs even to BMO, see [16].

Now, let us see how Theorem 1.1 improves the previously known estimates via nonlinear potentials. In [25] a Wolff potential type gradient bound has been obtained for equations without coefficients of the type $u_t - \operatorname{div} a(Du) = \mu$. More precisely, in [25] we introduced the following *intrinsic Wolff potentials*:

$$\mathbf{W}^{\mu}_{\lambda}(x_0,t_0;r) := \int_0^r \left(\frac{|\mu|(Q^{\lambda}_{\varrho}(x_0,t_0))}{\lambda^{2-p}\varrho^{N-1}}\right)^{1/(p-1)} \frac{d\varrho}{\varrho}\,, \qquad N = n+2 \qquad \lambda > 0\,.$$

See [14, 15, 18] for more on Wolff potentials. We then proved the existence of a universal constant $c_w \equiv c_w(n,p,\nu,L)$ for which

$$c_w \mathbf{W}^{\mu}_{\lambda}(x_0,t_0;r) + c_w \left(\fint_{Q^{\lambda}_r(x_0,t_0)} (|Du|+s)^{p-1}\,dx\,dt\right)^{1/(p-1)} \leq \lambda \\ \Longrightarrow |Du(x_0,t_0)| \leq \lambda \tag{1.24}$$

holds. Let us now show that (1.19) implies (1.24), for

$$c_w := \left[4^{N+1}(\log 2)^{2-p}c\right]^{1/(p-1)} + 4^N c \tag{1.25}$$

and radii $r \leq R_0$, where c and R_0 are the constants appearing in the statement of Theorem 1.1 (no restriction on radii appears in the case of equations as in (1.20)). With $r_i = r/2^i$ for integers $i \geq 0$, Hölder's inequality for series (as it is $p \geq 2$ here), gives

$$\begin{aligned}
\mathbf{I}^{\mu}_{1,\lambda}(x_0,t_0;r/2) &= \sum_{i=1}^{\infty}\int_{r_{i+1}}^{r_i} \frac{|\mu|(Q^{\lambda}_{\varrho}(x_0,t_0))}{\varrho^{N-1}}\frac{d\varrho}{\varrho} \\
&\leq 2^{N-1}\log 2 \sum_{i=1}^{\infty} \frac{|\mu|(Q^{\lambda}_{r_i}(x_0,t_0))}{r_i^{N-1}} \\
&\leq 2^{N-1}\log 2 \left[\sum_{i=1}^{\infty}\left(\frac{|\mu|(Q^{\lambda}_{r_i}(x_0,t_0))}{r_i^{N-1}}\right)^{1/(p-1)}\right]^{p-1} \\
&\leq 4^N(\log 2)^{2-p}\left[\sum_{i=1}^{\infty}\int_{r_i}^{r_{i-1}}\left(\frac{|\mu|(Q^{\lambda}_{\varrho}(x_0,t_0))}{\varrho^{N-1}}\right)^{1/(p-1)}\frac{d\varrho}{\varrho}\right]^{p-1} \\
&= 4^N(\log 2)^{2-p}\lambda^{2-p}\left[\mathbf{W}^{\mu}_{\lambda}(x_0,t_0;r)\right]^{p-1} \\
&\leq 4^N(\log 2)^{2-p}c_w^{1-p}\lambda \leq \frac{\lambda}{2c}\,.
\end{aligned}$$

Notice that to derive the second-last estimate we have use the inequality in the first line of (1.24), while in the last estimate we have used (1.25). Using the inequality

in the last display, again (1.25) and finally the left hand hand side of (1.24), a standard manipulation gives

$$c\mathbf{I}^{\mu}_{1,\lambda}(x_0,t_0;r/2)+c\left(\fint_{Q^{\lambda}_{r/2}(x_0,t_0)}(|Du|+s)^{p-1}\,dx\,dt\right)^{1/(p-1)}\leq\lambda$$

so that the right hand side inequality in (1.24), that is $|Du(x_0,t_0)|\leq\lambda$, follows applying Theorem 1.1.

The improvement from (1.24) to (1.19) is rather strong both from the viewpoint of the theoretical significance - as now the regularity theory of quasilinear equations is unified with that of the heat equation up to spatial gradient continuity - and from the one of the consequences. For instance, when looking for sharp criteria for establishing $Du\in L^{\infty}$ (and eventually in C^0) in terms of Lorentz spaces, the result in (1.24) gives that

$$\mu\in L(N,1/(p-1))\Longrightarrow Du\in L^{\infty}\,, \tag{1.26}$$

where we recall that

$$\mu\in L(N,1/(p-1))\Longleftrightarrow\int_0^{\infty}\lambda^{\frac{2-p}{p-1}}|\{(x,t)\in\Omega_T\,:\,|\mu(x,t)|>\lambda\}|^{\frac{1}{N(p-1)}}\,d\lambda<\infty\,.$$

The condition of Corollary 1.2 is clearly stronger that the one in (1.26), as $L(N,1)\subset L(N,1/(p-1))$, this inclusion being strict for $p>2$. For further properties of Lorentz spaces we refer for instance to [44]. Moreover, it is easy too see that more refined criteria in terms of density/concentration are provided by (1.19) with respect to (1.24) when μ is genuinely a measure. We also remark that Wolff potentials play a major role in the analysis of the fine properties of quasilinear equations (see for instance [17, 18, 39, 40, 45]); since the estimates contained in this paper are stronger than those involving Wolff potentials, we expect they will have a similar, if not stronger, impact in the future.

1.4. **Techniques.** Finally, a few comments on the methods used in this paper. Several new ingredients are needed to deal with the parabolic case with respect to the previous elliptic one [24], and the proofs depart considerably from those proposed before. The proof of Theorem 1.1 involves a very delicate, double-step induction procedure based on a few ingredients that re-shuffle, in a pointwise manner, some classical methods used in linear Calderón-Zygmund theory and combine them with the use of intrinsic geometry. Extensive use of nonlinear potential theoretic methods and regularity theory is made throughout. Let us briefly describe the heuristic used here by specializing for simplicity to the model case (1.1) and considering $\mu\in L^1$; the essence relies in careful procedure that allows to "linearize the equation" and control the possible degeneracy in a precisely quantified way at every scale. The following argument will be purely formal. We consider a dyadic shirking sequence of intrinsic cylinders for $i\geq 0$

$$\ldots Q^{\lambda}_{r_{i+1}}(x_0,t_0)\subset Q^{\lambda}_{r_i}(x_0,t_0)\subset Q^{\lambda}_{r_{i-1}}(x_0,t_0)\ldots\qquad r_i:=\sigma^i r$$

where $\sigma\in(0,1)$ is a constant depending only on n,p and λ is as in (1.19). A suitable exit time argument, together with very careful regularity estimates for solutions to homogeneous equations, gives

$$\lambda-\text{“quantified error”}\lesssim|Du|\quad\text{on } Q^{\lambda}_{r_i}(x_0,t_0)\ \text{ for } i \text{ large enough}\,. \tag{1.27}$$

This is something that in a way we can always assume, starting from an exit time index, otherwise we are going to get an opposite inequality for the integral averages $|(Du)_{Q^{\lambda}_{r_i}(x_0,t_0)}|\lesssim\lambda$, that eventually leads to an immediate proof of $|Du(x_0,t_0)|\leq$

λ, and therefore of (1.19). Assuming (1.27) leads to implement a delicate iteration procedure whose finally outcome is the following inequality:

$$\fint_{Q^\lambda_{r_j}(x_0,t_0)} |Du|^{p-1}\,dx\,dt \lesssim \lambda^{p-1} \tag{1.28}$$

that again implies (1.19). Note that proving (1.28) not only allows to implement the iteration but also allows to use, at each scale, the intrinsic geometry methods (compare with (1.14)). As emphasized, the key to the proof of Theorem 1.1 is the lower bound in (1.27); let us now give a formal but yet convincing argument showing how a condition as (1.27) allows to get (1.19) and why intrinsic Riesz potentials and and conditions as (1.19) naturally occur. Let us consider then (1.27) to be satisfied on $Q^\lambda_r(x_0,t_0)$ with a null error, i.e., $\lambda \lesssim |Du|$, and let us assume the first inequality in (1.19). The lower bound $\lambda \lesssim |Du|$ in turn allows to gain coercivity enough to treat the equation in display (1.1) as a heat equation with a coefficient, that is $u_t - \operatorname{div}(\lambda^{p-2}Du) = \mu$ that we can rewrite as

$$\lambda^{2-p}u_t - \triangle u = \lambda^{2-p}\mu \qquad \text{in } Q^\lambda_r(x_0,t_0)\,.$$

Now the effect of the use of intrinsic geometry and of the intrinsic Riesz potential shows-up. Changing variables and introducing

$$v(x,t) := \frac{u(x_0+rx, t_0+\lambda^{2-p}r^2t)}{r}\,, \quad \tilde\mu(x,t) := \lambda^{2-p}r\mu(x_0+rx,t_0+\lambda^{2-p}r^2t)$$

for $(x,t) \in Q_1 = B_1 \times (-1,0)$, we have

$$v_t - \triangle v = \tilde\mu\,. \tag{1.29}$$

Next, we apply the standard Riesz potential bound for solutions to (1.29), that is

$$|Dv(0,0)| \lesssim \mathbf{I}^{|\tilde\mu|}_1(0,0;1) + \left(\fint_{Q_1(0,0)} |Dv|^{p-1}\,dx\,dt\right)^{1/(p-1)}\,. \tag{1.30}$$

Changing variables back to μ we notice

$$\begin{aligned}
\mathbf{I}^{|\tilde\mu|}_1(0,0;1) &= \lambda^{2-p}r\int_0^1 \fint_{Q_\varrho(0,0)} |\mu(x_0+rx,t_0+\lambda^{2-p}r^2t)|\,dx\,dt\,d\varrho \\
&= \lambda^{2-p}r\int_0^1 \fint_{Q^\lambda_{\varrho r}(x_0,t_0)} |\mu(x,t)|\,dx\,dt\,d\varrho \\
&= \lambda^{2-p}\int_0^r \fint_{Q^\lambda_{\varrho}(x_0,t_0)} |\mu(x,t)|\,dx\,dt\,d\varrho = \lambda^{2-p}\mathbf{I}^{\mu}_{1,\lambda}(x_0,t_0;r) \lesssim \lambda\,,
\end{aligned}$$

where in the last inequality we have used the first line in (1.19). Finally, scaling back to u, using the previous inequality together with the first line of (1.19), we conclude with $|Du(x_0,t_0)| = |Dv(0,0)| \lesssim \lambda$, that is the proof of (1.19). The one outlined in the last lines is only a heuristic argument used to show how intrinsic Riesz potentials play a decisive and natural role in this context, but its rigorous implementation is highly nontrivial and involves a refined double induction argument that exploits rather subtle aspects of regularity theory of degenerate parabolic equations. Several tools are used here. One of the main points is that the analysis of the relevant iterating quantities must be performed at two different levels, using different energy spaces. Indeed, since we are dealing essentially with measure data problems, the natural spaces involved are larger than $L^p(-T,0;W^{1,p})$. This, together with the lack of reverse Hölder type inequalities and homogeneous estimates which is typical when dealing with degenerate parabolic equations, reflects in a simultaneous use of two different spaces, namely $L^1(-T,0;W^{1,1})$ and $L^{p-1}(-T,0;W^{1,p-1})$. Eventually, a very delicate interplay between local regularity

of solutions to homogeneous equations and comparison estimates must be exploited in the framework of intrinsic geometries thanks to exit time arguments and the use of intrinsic Riesz potentials. The proof of Theorem 1.1 eventually opens the way to the continuity analysis and in particular to Theorem 1.3. For this we shall readapt the iteration procedure of Theorem 1.1 to estimate oscillations rather than the size of the gradient. This in turn imposes to consider a priori infinite many exit times arguments used to control the degeneracy of the equation via the oscillations of the gradient, and vice-versa.

2. Preparations

2.1. **General notation.** In what follows we denote by c a general positive constant, possibly varying from line to line; special occurrences will be denoted by $c_1, c_2, \bar{c}_1, \bar{c}_2$ or the like. All these constants will always be larger or equal than one; moreover relevant dependencies on parameters will be emphasized using parentheses, i.e., $c_1 \equiv c_1(n,p,\nu,L)$ means that c_1 depends only on n,p,ν,L. We denote by $B(x_0,r) := \{x \in \mathbb{R}^n \, : \, |x-x_0| < r\}$ the open ball with center x_0 and radius $r>0$; when not important, or clear from the context, we shall omit denoting the center as follows: $B_r \equiv B(x_0,r)$. Unless otherwise stated, different balls in the same context will have the same center. We shall also denote $B \equiv B_1 = B(0,1)$ if not differently specified. In a similar fashion standard and intrinsic parabolic cylinders with vertex (x_0,t_0) and width $r>0$ have been defined in (1.12) and (1.13), respectively. When the vertex will not be important in the context or it will be clear that all the cylinders occurring in a proof will share the same vertex, we shall omit to indicate it, simply denoting Q_r and Q_r^λ for the cylinders in (1.12) and (1.13), respectively. We recall that if $Q = \mathcal{A} \times (t_1,t_2)$ is a cylindrical domain, the usual parabolic boundary of Q is $\partial_{\rm par} Q := (\mathcal{A} \times \{t_1\}) \cup (\partial\mathcal{A} \times [t_1,t_2))$, and this is nothing else but the standard topological boundary without the upper cap $\bar{\mathcal{A}} \times \{t_2\}$. Accordingly, we define the parabolic closure of Q as

$$\lfloor Q \rfloor_{\rm par} := Q \cup \partial_{\rm par} Q \,. \tag{2.1}$$

With $\mathcal{O} \subset \mathbb{R}^{n+1}$ being a measurable subset with positive measure, and with $g \colon \mathcal{O} \to \mathbb{R}^n$ being a measurable map, we shall denote by

$$(g)_{\mathcal{O}} \equiv \fint_{\mathcal{O}} g(x,t)\, dx\, dt := \frac{1}{|\mathcal{O}|} \int_{\mathcal{O}} g(x,t)\, dx\, dt$$

its integral average; here $|\mathcal{O}|$ denotes the Lebesgue measure of $\mathcal{O}$. A similar notation is adopted if the integral is only in space or time. In the rest of the paper we shall use several times the following elementary property of integral averages:

$$\left(\fint_{\mathcal{O}} |g-(g)_{\mathcal{O}}|^q \, dx\, dt \right)^{1/q} \le 2 \left(\fint_{\mathcal{O}} |g-\gamma|^q \, dx\, dt \right)^{1/q} , \tag{2.2}$$

whenever $\gamma \in \mathbb{R}^n$ and $q \ge 1$. The oscillation of g on $\mathcal{O}$ is instead defined as

$$\operatorname*{osc}_{\mathcal{O}} g := \sup_{(x,t),(\tilde{x},\tilde{t}) \in \mathcal{O}} |g(x,t) - g(\tilde{x},\tilde{t})| \,.$$

Finally, we remark that we shall denote the partial derivative with respect to time of a function u both by u_t and by $\partial_t u$; moreover, the letter λ will always denote a positive number. Further relevant notation is at the beginning of the next section.

2.2. **Comparison maps.** The basic setup in this section is tailored to the needs of the proof of Theorem 1.1 and subsequent results. Therefore we shall consider u to be an energy solution to (1.3) under the assumptions (1.4)-(1.5) and (1.8) until the end Section 2.4; only in Section 2.5 we shall discuss the general case, thereby

treating SOLA and discarding assumption (1.8). With a point $(x_0, t_0) \in \Omega_T$ being fixed, and given an intrinsic cylinder of the type

$$Q_r^\lambda(x_0, t_0) \equiv B(x_0, r) \times (t_0 - \lambda^{2-p} r^2, t_0)$$

such that $Q_{2r}^\lambda(x_0, t_0) \subset \Omega_T$, we consider a family of nested parabolic cylinders

$$Q_j \equiv B_j \times T_j \equiv B(x_0, r_j) \times (t_0 - \lambda^{2-p} r_j^2, t_0) \subseteq \Omega_T \,, \qquad r_j := \sigma^j r \,, \tag{2.3}$$

for a fixed decay parameter $\sigma \in (0, 1/4)$; notice that we always have

$$\ldots Q_j \subset \frac{1}{4} Q_{j-1} \subset Q_{j-1} \ldots$$

Accordingly, we consider their dyadic, parabolic dilations

$$\tau Q_j \equiv Q_{\tau r_j}^\lambda(x_0, t_0) \equiv \tau B_j \times \tau T_j \equiv B(x_0, \tau r_j) \times (t_0 - \lambda^{2-p} (\tau r_j)^2, t_0)$$

for $\tau > 0$; notice that here, slightly abusing the notation, we are denoting

$$\tau T_j \equiv (t_0 - \lambda^{2-p} (\tau r_j)^2, t_0) \,.$$

A similar notation will occur several times in rest of the paper. Now, let

$$w_j \in C^0(T_j; L^2(B_j)) \cap L^p(T_j; W^{1,p}(B_j))$$

be the unique solution to the Cauchy-Dirichlet problem

$$\begin{cases} \partial_t w_j - \operatorname{div} a(x, t, Dw_j) = 0 & \text{in } Q_j \\ w_j = u & \text{on } \partial_{\text{par}} Q_j \,. \end{cases} \tag{2.4}$$

After having defined w_j, we also define

$$v_j \in C^0(\tfrac{1}{2} T_j; L^2(\tfrac{1}{2} B_j)) \cap L^p(\tfrac{1}{2} T_j; W^{1,p}(\tfrac{1}{2} B_j))$$

as the unique solution to the frozen Cauchy-Dirichlet problem

$$\begin{cases} \partial_t v_j - \operatorname{div} a(x_0, t, Dv_j) = 0 & \text{in } \frac{1}{2} Q_j \\ v_j = w_j & \text{on } \partial_{\text{par}} \left(\frac{1}{2} Q_j\right) . \end{cases} \tag{2.5}$$

2.3. **A priori estimates for comparison maps.** We now derive various a priori estimates for w_j and v_j, starting from L^∞-bounds. When turning our attention to w_j we need to use the results recently established in [29], that allow to deal with equations with non-constant, Dini-continuous spatial coefficients. We start with a statement in terms of intrinsic geometry.

Theorem 2.1 (Intrinsic gradient bound)**.** *Let w_j be as in* (2.4)*. There exists a positive radius $R_1 \equiv R_1(n, p, \nu, L, \omega(\cdot))$ and a constant $c_1 \equiv c_1(n, p, \nu, L)$ such that if $\varrho \in (0, R_1)$ and*

$$Q_\varrho^{\lambda_0}(x_1, t_1) := B(x_1, \varrho) \times (t_1 - \lambda_0^{2-p} \varrho^2, t_1) \subset Q_j \,,$$

then the implication

$$c_1 \left(\fint_{Q_\varrho^{\lambda_0}(x_1, t_1)} (|Dw_j| + s)^p \, dx \, dt \right)^{1/p} \leq \lambda_0 \quad \Longrightarrow \quad |Dw_j(x_1, t_1)| \leq \lambda_0$$

holds.

Proof. This result has been proved and used in [29, Theorem 1.1, Theorem 1.3, Theorem 4.1]. The proof is exactly the one given in the proof of [29, Theorem 1.1], once [29, Lemma 4.3] is used instead of [29, Lemma 4.2], see also [29, Remark 4.1]. We also remark that no restriction on ϱ is necessary when the vector field $a(\cdot)$ is independent of x. □

The pointwise bound of Theorem 2.1 can be turned into an L^∞-bound of exactly the same type proved by DiBenedetto [9] for equations with no coefficients, see also Theorem 2.2 below. Since we are going to cover the case of equations with measure data, where solutions with low degree of integrability naturally appear, we need to lower the p-integrability exponent to $(p-1)$ to get the correct form of a priori estimates. All this is done in the next

Corollary 2.1. *Let w_j be as in (2.4) and R_1 as in Theorem 2.1. There exists a constant $c_2 \equiv c_2(n,p,\nu,L)$ such that if $r \in (0,R_1)$, then*

$$\sup_{\tau_m Q_j} |Dw_j| + s \leq c_2(\lambda+s) + \frac{c_2\lambda^{2-p}}{2^{n+2}(1-\tau_m)^{n+2}} \fint_{Q_j} (|Dw_j|+s)^{p-1}\,dx\,dt$$

holds whenever $\tau_m \in (0,1)$. In particular, we have

$$\sup_{\frac{1}{2}Q_j} |Dw_j| + s \leq c_2(\lambda+s) + c_2\lambda^{2-p} \fint_{Q_j} (|Dw_j|+s)^{p-1}\,dx\,dt\,.$$

Proof. Define

$$\lambda_0 \equiv \lambda_0(\tau,\tau') := \frac{1}{2}\sup_{\tau Q_j} |Dw_j| + \lambda + s + \frac{2c_1^p}{(\tau-\tau')^{n+2}}\lambda^{2-p} \fint_{Q_j} (|Dw_j|+s)^{p-1}\,dx\,dt$$

whenever $\tau_m \leq \tau' < \tau \leq 1$, where $c_1 \equiv c_1(n,p,\nu,L)$ is as in Theorem 2.1. As $\lambda_0 \geq \lambda$ and $p \geq 2$, we clearly have that for $\delta := \tau - \tau'$ the inclusion

$$Q^{\lambda_0}_{\delta r_j}(x_1,t_1) := B(x_1,\delta r_j) \times (t_1 - \lambda_0^{2-p}(\delta r_j)^2, t_1) \subset \tau Q_j = \tau Q^{\lambda}_{r_j}(x_0,t_0)$$

holds whenever $(x_1,t_1) \in \tau' Q_j$. Furthermore, by using the very definition of λ_0 we may estimate

$$\begin{aligned}
& c_1 \left(\fint_{Q^{\lambda_0}_{\delta r_j}(x_1,t_1)} (|Dw_j|+s)^p\,dx\,dt \right)^{1/p} \\
&\quad \leq c_1 2^{1/p}\lambda_0^{1/p} \left(\fint_{Q^{\lambda_0}_{\delta r_j}(x_1,t_1)} (|Dw_j|+s)^{p-1}\,dx\,dt \right)^{1/p} \\
&\quad \leq c_1 2^{1/p}\lambda_0^{1/p} \left(\frac{|Q_j|}{|Q^{\lambda_0}_{\delta r_j}(x_1,t_1)|} \right)^{1/p} \left(\fint_{Q_j} (|Dw_j|+s)^{p-1}\,dx\,dt \right)^{1/p} \\
&\quad = c_1 2^{1/p}\lambda_0^{1/p} \left(\frac{\lambda^{2-p}}{\delta^{n+2}\lambda_0^{2-p}} \right)^{1/p} \left(\fint_{Q_j} (|Dw_j|+s)^{p-1}\,dx\,dt \right)^{1/p} \\
&\quad \leq c_1 2^{1/p}\lambda_0^{1/p} \left(\frac{\lambda^{2-p}}{\delta^{n+2}\lambda_0^{2-p}} \right)^{1/p} \left(\frac{\lambda_0\delta^{n+2}}{2c_1^p\lambda^{2-p}} \right)^{1/p} = \lambda_0\,.
\end{aligned}$$

Therefore Theorem 2.1 implies that $|Dw_j(x_1,t_1)| \leq \lambda_0$. But this holds for all $(x_1,t_1) \in \tau' Q_j$ and thus

$$\sup_{\tau' Q_j} |Dw_j| \leq \frac{1}{2}\sup_{\tau Q_j} |Dw_j| + \lambda + s + \frac{2c_1^p\lambda^{2-p}}{(\tau-\tau')^{n+2}} \fint_{Q_j} (|Dw_j|+s)^{p-1}\,dx\,dt$$

follows. Lemma 2.1 below applied with $\varphi(\tau) = \sup_{\tau Q_j} |Dw_j|$ then concludes the proof by properly choosing the constant c_2. □

The nest result is a classical iteration lemma for the of which we refer to [31, Lemma 6.1].

Lemma 2.1. *Let $\varphi : [\tau_m, 1] \to [0, \infty)$, with $\tau_m \in (0,1)$, be a function such that*

$$\varphi(\tau') \leq \frac{1}{2}\varphi(\tau) + K + \frac{\mathcal{B}}{(\tau - \tau')^{n+2}} \quad \text{holds for every} \quad \tau_m \leq \tau' < \tau \leq 1 \,,$$

where $\mathcal{B}, K \geq 0$. Then $\varphi(\tau_m) \leq c(n)K + (1-\tau_m)^{-(n+2)}\mathcal{B}$.

Corollary 2.1 obviously holds for v_j too, and in this case it is a by now classical estimate of DiBenedetto [9], as already mentioned above. See also for example [1, 25] for similar bounds. We report the statement for completeness.

Theorem 2.2. *Let v_j be as in* (2.5)*. For a constant $c_3 \equiv c_3(n,p,\nu,L)$ we have*

$$\sup_{\frac{1}{4}Q_j} |Dv_j| + s \leq c_3(\lambda + s) + c_3\lambda^{2-p} \fint_{\frac{1}{2}Q_j} (|Dv_j| + s)^{p-1}\, dx\, dt \,.$$

We now pass to give oscillation estimates for w_j and v_j. The next result provides a gradient oscillations estimate for solutions to homogeneous equations with Dini-continuous coefficients.

Theorem 2.3 (Continuity estimate)**.** *Let w_j be as in* (2.4)*, then Dw_j is continuous. Moreover, assume that*

$$\sup_{\frac{1}{2}Q_j} |Dw_j| + s \leq A\lambda \tag{2.6}$$

holds for some $A \geq 1$. Then, for any $\delta \in (0,1)$ there exists a positive constant $\sigma_1 \equiv \sigma_1(n,p,\nu,L,A,\delta,\omega(\cdot)) \in (0,1/4)$ such that

$$\operatorname*{osc}_{\sigma_1 Q_j} Dw_j \leq \delta\lambda \,. \tag{2.7}$$

Proof. The starting point of the proof is the work in [29], where the continuity of the gradient of solutions to parabolic equations as in (1.3) has been proved under the assumption of Dini-continuity of the space coefficients; this by the way immediately implies the continuity of Dw_j claimed in the statement. What we need here is a quantitative bound on the oscillations of Dw_j. To this aim, let us briefly recall the main arguments in [29, proof of Theorem 1.3], where the continuity properties of Du are formulated and proved in terms of the auxiliary vector field

$$V(z) := (|z|^2 + s^2)^{(p-2)/4} z$$

and the related field $V(Dw_j)$. For the use of such maps in the present context we refer to [29]; the only property we shall use here is the following inequality:

$$|z_1 - z_2| \leq c_v \frac{|V(z_1) - V(z_2)|}{(s^2 + |z_1|^2 + |z_2|^2)^{(p-2)/4}} \,, \tag{2.8}$$

that holds for $c_v \equiv c_v(n,p)$ and for all vectors $z_1, z_2 \in \mathbb{R}^n$ which are not simultaneously null; see for instance [37]. By following the arguments developed for [29, (5.15)] it can be proved that for every $\varepsilon \in (0,1)$ there exists a positive radius $R_\varepsilon \equiv R_\varepsilon(n,p,\nu,L,\omega(\cdot),\varepsilon) \in (0,1/16)$ such that

$$|(V(Dw_j))_{Q_\tau^{A\lambda}(\tilde{x},\tilde{t})} - (V(Dw_j))_{Q_\varrho^{A\lambda}(\tilde{x},\tilde{t})}| \leq (A\lambda)^{p/2}\varepsilon \tag{2.9}$$

and

$$\left(\fint_{Q_\varrho^{A\lambda}(\tilde{x},\tilde{t})} |V(Dw_j) - (V(Dw_j))_{Q_\varrho^{A\lambda}(\tilde{x},\tilde{t})}|^2 \, dx\, dt \right)^{1/2} \leq (A\lambda)^{p/2}\varepsilon \tag{2.10}$$

hold whenever $(\tilde{x},\tilde{t}) \in \frac{1}{4}Q_j$ and $0 < \tau \leq \varrho \leq R_\varepsilon r_j$; notice that R_ε is in particular independent of λ, A and the considered cylinder Q_j. Letting $\tau \to 0$ in (2.9) and recalling that $V(Dw_j)$ is continuous yields

$$|V(Dw_j(\tilde{x},\tilde{t})) - (V(Dw_j))_{Q_\varrho^{A\lambda}(\tilde{x},\tilde{t})}| \leq (A\lambda)^{p/2}\varepsilon \qquad \forall\ \varrho \in (0, R_\varepsilon r_j] \,. \tag{2.11}$$

We are now ready to finish the proof with the choice

$$\varepsilon := \frac{\delta^{p/2}}{c_v 2^{p/2-1}48^N A^{p/2}}, \qquad \sigma_1 := \frac{A^{(2-p)/2}R_\varepsilon}{32}\,.$$

The constant c_v is the one appearing in (2.8) Notice that the dependence of σ_1 upon $n, p, \nu, L, A, \delta, \omega(\cdot)$, as described in the statement, appears through the one implicitly contained in R_ε. Take now $(\tilde y, \tilde s), (\tilde x, \tilde t) \in \sigma_1 Q_j$; we can assume that $\tilde t \geq \tilde s$ otherwise we can exchange the role between the two points in the next lines. It obviously follows that

$$Q^{A\lambda}_{R_\varepsilon r_j/8}(\tilde y, \tilde s) \subset Q^{A\lambda}_{R_\varepsilon r_j}(\tilde x, \tilde t) \subset \frac{1}{4}Q_j\,.$$

Using this last fact, thanks to Jensen's inequality, the one in display (2.10) and using also (2.2), we have

$$\begin{aligned}
&|(V(Dw_j))_{Q^{A\lambda}_{R_\varepsilon r_j/8}(\tilde y,\tilde s)} - (V(Dw_j))_{Q^{A\lambda}_{R_\varepsilon r_j/8}(\tilde x,\tilde t)}| \\
&\leq \fint_{Q^{A\lambda}_{R_\varepsilon r_j/8}(\tilde y,\tilde s)} |V(Dw_j) - (V(Dw_j))_{Q^{A\lambda}_{R_\varepsilon r_j/8}(\tilde x,\tilde t)}|\,dx\,dt \\
&\leq 2\fint_{Q^{A\lambda}_{R_\varepsilon r_j/8}(\tilde y,\tilde s)} |V(Dw_j) - (V(Dw_j))_{Q^{A\lambda}_{R_\varepsilon r_j}(\tilde x,\tilde t)}|\,dx\,dt \\
&\leq 16^N\fint_{Q^{A\lambda}_{R_\varepsilon r_j}(\tilde x,\tilde t)} |V(Dw_j) - (V(Dw_j))_{Q^{A\lambda}_{R_\varepsilon r_j}(\tilde x,\tilde t)}|\,dx\,dt \\
&\leq 16^N\left(\fint_{Q^{A\lambda}_{R_\varepsilon r_j}(\tilde x,\tilde t)} |V(Dw_j) - (V(Dw_j))_{Q^{A\lambda}_{R_\varepsilon r_j}(\tilde x,\tilde t)}|^2\,dx\,dt\right)^{1/2} \\
&\leq 16^N (A\lambda)^{p/2}\varepsilon\,.
\end{aligned}$$

By using the previous estimate and (2.11) (actually used also for $(\tilde y, \tilde s)$ instead of $(\tilde x, \tilde t)$) together with triangle inequality we easily gain

$$|V(Dw_j(\tilde x,\tilde t)) - V(Dw_j(\tilde y,\tilde s))| \leq 48^N(A\lambda)^{p/2}\varepsilon\,. \tag{2.12}$$

We are now ready to show (2.7) proving that

$$|Dw_j(\tilde x,\tilde t) - Dw_j(\tilde y,\tilde s)| \leq \delta\lambda \tag{2.13}$$

whenever $(\tilde y, \tilde s), (\tilde x, \tilde t) \in \sigma_1 Q_j$. First of all, observe that we can assume that either $|Dw_j(\tilde x,\tilde t)| \geq \delta\lambda/2$ or $|Dw_j(\tilde y,\tilde s)| \geq \delta\lambda/2$ holds otherwise we are done. The inequalities in (2.8) and (2.12) then imply (2.13) as follows:

$$|Dw_j(\tilde x,\tilde t) - Dw_j(\tilde y,\tilde s)| \leq c_v 48^N(A\lambda)^{p/2}(\delta\lambda/2)^{1-p/2}\varepsilon = \delta\lambda$$

and the proof is complete. □

Remark 2.1. The proof of the previous result allows in fact, together with the argument given in [29], to get an explicit modulus of continuity for the gradient of solutions of equations with Dini-continuous coefficients. Indeed, the choice of the radius $R_\varepsilon r_j$ making (2.9) is made to meet a condition of the form

$$\omega(R_\varepsilon r_j) + \int_0^{R_\varepsilon r_j} \omega(\varrho)\,\frac{d\varrho}{\varrho} \leq \frac{\varepsilon^b}{c}$$

for some positive constants b and c depending only on n, p, ν, L, A. This gives a modulus of continuity involving a power of the function

$$r \mapsto \omega(r) + \int_0^r \omega(\varrho)\,\frac{d\varrho}{\varrho}$$

which is in accordance to the known results in the classical elliptic regularity theory.

We now collect a few results from [25, 26] in order to provide oscillation estimates for the functions v_j. The next statement is a slight variant of [25, Theorem 3.1].

Theorem 2.4. *Let v_j be as in* (2.5). *Consider numbers*

$$A, B \geq 1 \qquad \text{and} \qquad \bar{\varepsilon} \in (0,1)\,.$$

Then there exists a constant $\sigma_2 \in (0, 1/4)$ depending only on $n, p, \nu, L, A, B, \bar{\varepsilon}$ such that if

$$\frac{\lambda}{B} \leq \sup_{\sigma_2 Q_j} |Dv_j| \leq s + \sup_{\frac{1}{4}Q_j} |Dv_j| \leq A\lambda \tag{2.14}$$

holds, then

$$\fint_{\tau Q_j} |Dv_j - (Dv_j)_{\tau Q_j}|\, dx\, dt \leq \bar{\varepsilon} \fint_{\frac{1}{4}Q_j} |Dv_j - (Dv_j)_{\frac{1}{4}Q_j}|\, dx\, dt \tag{2.15}$$

holds too, whenever $\tau \in (0, \sigma_2]$.

Remark 2.2. The essence of the previous result lies in the fact that once the bounds (2.14) are satisfied, then solutions to evolutionary p-Laplacean type equations satisfy elliptic type decay estimates as in (2.17) when framed in the proper intrinsic geometry dictated by (2.14). Indeed, let us denote by $E(f, Q)$ the usual excess functional

$$E(f, Q) := \fint_Q |f - (f)_Q|\, dx\, dt \tag{2.16}$$

which is defined whenever f is an integrable function and Q a measurable set with positive measure; this functional gives an integral measure of the oscillations of f in a subset Q. Estimate (2.15) now reads as

$$E(Dv_j, \tau Q_j) \leq \bar{\varepsilon} E(Dv_j, \tfrac{1}{4}Q_j)\,. \tag{2.17}$$

Theorem 2.4 gives the natural analog, when passing to the framework of degenerate parabolic equations of p-Laplacean type, of the classical results known for solutions to the heat equations. Indeed, Theorem 2.4 holds without assuming (2.14) for solutions to (1.2). This is a classical result of Campanato [5].

Using Theorem 2.4 it is possible to give a proof of the Hölder continuity of the gradient of solutions to frozen equations, as for instance shown in [26, Theorem 3.2]; see also [25, Theorem 3.2].

Theorem 2.5. *Let v_j be as in* (2.5). *For every $A \geq 1$ there exist constants $c_4 \equiv c_4(n, p, \nu, L, A)$ and $\alpha \equiv \alpha(n, p, \nu, L, A)$ such that*

$$\sup_{\frac{1}{4}Q_j} |Dv_j| + s \leq A\lambda \qquad \Longrightarrow \qquad \operatorname*{osc}_{\tau Q_j} Dv_j \leq c_4 \tau^{\alpha} \lambda \qquad \forall\, \tau \in (0, 1/4)\,.$$

Remark 2.3. Theorems 2.4-2.5 have been presented in [25, 26] actually for solutions to equations of the type

$$u_t - \operatorname{div} a(Du) = 0\,. \tag{2.18}$$

On the other hand, by following the arguments in [25, 26] it is not difficult to see that all the proofs carry out for solutions to equations of the type in display (1.20). Therefore Theorems 2.4-2.5 apply to the functions v_j as well, that indeed solve equations as in (1.20).

2.4. **Comparison estimates.** We start this section by a reformulation of a result established in [25, Lemma 4.1] and [25, (4.5), (4.6)]. We remark that the result there was presented only for equations without coefficients as in (2.18). Nevertheless, the proof works directly for general equations with merely measurable coefficients; the crucial point is the strict monotonicity in the gradient variable.

Lemma 2.2. *Let u be as in Theorem 1.1 and w_j as in (2.4) with $j \geq 0$. Let $\tilde{\varepsilon} \in (0, 1/(n+1)]$. There exist constants $\bar{c}_1 \equiv \bar{c}_1(n,p,\nu,\tilde{\varepsilon})$ and $\bar{c}_2 \equiv \bar{c}_2(n,p,\nu)$ such that*

$$\left(\fint_{Q_j} |Du - Dw_j|^q \, dx \, dt \right)^{1/q} \leq \bar{c}_1 \lambda \left[\frac{1}{\lambda} \frac{|\mu|(Q_j)}{r_j^{N-1}} \right]^{(n+2)/[(p-1)n+p]} \tag{2.19}$$

holds for any $0 < q \leq p - 1 + 1/(n+1) - \tilde{\varepsilon}$. Moreover, the inequalities

$$\sup_{t \in T_j} \int_{B_j} |u - w_j| \, dx \leq |\mu|(Q_j) \tag{2.20}$$

and

$$\fint_{Q_j} \frac{(|Du| + |Dw_j|)^{p-2} |Du - Dw_j|^2}{(\alpha + |u - w_j|)^{\xi}} \, dx \, dt \leq \bar{c}_2 \frac{\alpha^{1-\xi}}{\xi - 1} \left[\frac{|\mu|(Q_j)}{\lambda^{2-p} r_j^N} \right] \tag{2.21}$$

hold for any $\alpha > 0$ and $\xi > 1$.

We here recall a parabolic Sobolev-Poincaré inequality that will be useful in the sequel; we refer to [9, Chapter 1, Proposition 3.1] for the proof.

Proposition 2.1. *Let $v \in L^\infty(T_j; L^m(B_j)) \cap L^{q_2}(T_j; W_0^{1,q_2}(B_j))$ for $q_2, m \geq 1$. There exists a constant c depending only on n, q_2, m such that the following inequality holds for $q_1 = q_2(n+m)/n$:*

$$\int_{Q_j} |v|^{q_1} \, dx \, dt \leq c \left(\int_{Q_j} |Dv|^{q_2} \, dx \, dt \right) \left(\sup_{\tau} \int_{B_j} |v(x,\tau)|^m \, dx \right)^{q_2/n} .$$

Using the previous result and Lemma 2.2 we get another comparison estimate.

Lemma 2.3. *Let u be as in Theorem 1.1 and w_{j-1}, w_j as in (2.4), with $j \geq 1$. Then, for any $\tilde{\varepsilon} \in (0, 1/(n+1)]$, there exists a constant $\bar{c}_3 \equiv \bar{c}_3(n,p,\nu,\tilde{\varepsilon},\sigma)$ such that the inequality*

$$\left(\fint_{Q_j} |u - w_j|^q \, dx \, dt \right)^{1/q} \leq \bar{c}_3 r_{j-1} \lambda \left[\frac{1}{\lambda} \frac{|\mu|(Q_{j-1})}{r_{j-1}^{N-1}} \right]^{(n+p)/[(p-1)n+p]} \tag{2.22}$$

holds whenever $0 < q \leq p - 1 + p/n - \tilde{\varepsilon}$ and

$$\left(\fint_{Q_j} |Dw_{j-1} - Dw_j|^q \, dx \, dt \right)^{1/q} \leq \bar{c}_3 \lambda \left[\frac{1}{\lambda} \frac{|\mu|(Q_{j-1})}{r_{j-1}^{N-1}} \right]^{(n+2)/[(p-1)n+p]} \tag{2.23}$$

holds whenever $0 < q \leq p - 1 + 1/(n+1) - \tilde{\varepsilon}$.

Proof. To prove (2.22), we use Proposition 2.1 with the choice $m = 1$,

$$q_1 := p - 1 + \frac{p}{n} - \tilde{\varepsilon} , \qquad q_2 := \frac{n}{n+1} q_1 = p - 1 + \frac{1}{n+1} - \frac{n\tilde{\varepsilon}}{n+1}$$

and $v \equiv u - w_j \in L^\infty(T_j; L^2(B_j)) \cap L^{q_2}(T_j; W_0^{1,q_2}(B_j))$; recall (2.20). This yields

$$\left(\fint_{Q_j} |u - w_j|^{q_1} \, dx \, dt \right)^{1/q_1}$$

$$\leq c\left(\left[\fint_{Q_j} |Du - Dw_j|^{q_2}\,dx\,dt\right]^{1/q_2}\left[\sup_{\tau}\int_{B_j} |u - w_j|\,dx\right]^{1/n}\right)^{n/(n+1)}.$$

Let $\bar{c}_1$ be as in Lemma 2.2 be the constant corresponding to the choice $\tilde{\varepsilon}n/(n+1)$ instead of $\tilde{\varepsilon}$. Substituting (2.19) and (2.20) into the previous estimate leads to

$$\begin{aligned}
&\left(\fint_{Q_j} |u - w_j|^{q_1}\,dx\,dt\right)^{1/q_1}\\
&\quad\leq c\bar{c}_1^{n/(n+1)}\left(\lambda\left[\frac{1}{\lambda}\frac{|\mu|(Q_j)}{r_j^{N-1}}\right]^{(n+2)/[(p-1)n+p]}[|\mu|(Q_j)]^{1/n}\right)^{n/(n+1)}\\
&\quad\leq \bar{c}_3 r_{j-1}\lambda\left[\frac{1}{\lambda}\frac{|\mu|(Q_{j-1})}{r_{j-1}^{N-1}}\right]^{(n+p)/[(p-1)n+p]},
\end{aligned}$$

which, together with Hölder's inequality, proves (2.22). Here we have also used that $N-1 = n+1$ and the identity

$$\frac{n}{n+1}\left[\frac{n+2}{(p-1)n+p} + \frac{1}{n}\right] = \frac{n+p}{(p-1)n+p}.$$

As for (2.23) we instead argue as follows:

$$\begin{aligned}
&\left(\fint_{Q_j} |Dw_{j-1} - Dw_j|^{q_2}\,dx\,dt\right)^{1/q_2}\\
&\quad\leq \left(\frac{|Q_{j-1}|}{|Q_j|}\right)^{1/q_2}\left(\fint_{Q_{j-1}} |Du - Dw_{j-1}|^{q_2}\,dx\,dt\right)^{1/q_2}\\
&\qquad\qquad + \left(\fint_{Q_j} |Du - Dw_j|^{q_2}\,dx\,dt\right)^{1/q_2}\\
&\quad\leq \bar{c}_1\left[\left(\frac{|Q_{j-1}|}{|Q_j|}\right)^{1/q_2} + \frac{r_{j-1}^{N-1}}{r_j^{N-1}}\right]\lambda\left[\frac{1}{\lambda}\frac{|\mu|(Q_{j-1})}{r_{j-1}^{N-1}}\right]^{(n+2)/[(p-1)n+p]}\\
&\quad\leq \bar{c}_3\lambda\left[\frac{1}{\lambda}\frac{|\mu|(Q_{j-1})}{r_{j-1}^{N-1}}\right]^{(n+2)/[(p-1)n+p]},
\end{aligned}$$

where we have repeatedly applied (2.19) and used the fact that we are assuming $p \geq 2$. Now (2.23) follows, again by Hölder's inequality as $\tilde{\varepsilon}$ is arbitrary. □

The following lemma provides one of the key estimates to obtain Theorem 1.1.

Lemma 2.4. *Let u be as in Theorem 1.1 and w_{j-1}, w_j as in (2.4), with $j \geq 1$. Suppose further that*

$$\frac{|\mu|(Q_{j-1})}{r_{j-1}^{N-1}} \leq \lambda \tag{2.24}$$

and that the bounds

$$\frac{\lambda}{A} \leq |Dw_{j-1}| \leq A\lambda \quad \text{in } Q_j \tag{2.25}$$

hold for some $A \geq 1$. Then there exists a constant $\bar{c}_4$ depending only on n, p, ν, σ, A such that

$$\fint_{Q_j} |Du - Dw_j|\,dx\,dt \leq \bar{c}_4\left[\frac{|\mu|(Q_{j-1})}{r_{j-1}^{N-1}}\right]. \tag{2.26}$$

Proof. Let us begin by fixing several parameters appearing in the proof:

$$\gamma := \frac{1}{4(p-1)(n+1)}\,, \qquad \xi := 1+2\gamma\,, \qquad \tilde{\varepsilon} := \frac{1}{2(n+1)}\,, \tag{2.27}$$

and, throughout the proof, we will apply Lemmas 2.2 and 2.3 with exponents

$$0 < q \leq \bar{q} := \xi(p-1) = p-1+\frac{1}{2(n+1)}\,.$$

Let $\bar{c}_1$, $\bar{c}_2$, $\bar{c}_3$ be as in Lemma 2.2 and Lemma 2.3, respectively, corresponding to these choices of σ and $\tilde{\varepsilon}$; therefore they ultimately depend only on n, p, ν, σ. We also set

$$\bar{w}_{j-1} := \frac{w_{j-1}}{\lambda}\,, \qquad \bar{w}_j := \frac{w_j}{\lambda}\,. \tag{2.28}$$

In what follows constants denoted by c will only depend on n, p, ν, σ, A and will in general vary from line to line. We start to estimate the term on the left in (2.26) with the aid of (2.25) as follows:

$$\begin{aligned}
\fint_{Q_j} |Du - Dw_j|\,dx\,dt &\leq A^{(p-2)(1+\gamma)} \fint_{Q_j} |D\bar{w}_{j-1}|^{(p-2)(1+\gamma)} |Du - Dw_j|\,dx\,dt \\
&\leq c \fint_{Q_j} |D\bar{w}_j - D\bar{w}_{j-1}|^{(p-2)(1+\gamma)} |Du - Dw_j|\,dx\,dt \\
&\quad + c \fint_{Q_j} |D\bar{w}_j|^{(p-2)(1+\gamma)} |Du - Dw_j|\,dx\,dt\,.
\end{aligned} \tag{2.29}$$

Appealing to Hölder's inequality, together with (2.19) and (2.23), gives us

$$\begin{aligned}
&\fint_{Q_j} |D\bar{w}_j - D\bar{w}_{j-1}|^{(p-2)(1+\gamma)} |Du - Dw_j|\,dx\,dt \\
&\leq \lambda^{-(p-2)(1+\gamma)} \left(\fint_{Q_j} |Dw_j - Dw_{j-1}|^{(p-1)(1+\gamma)}\,dx\,dt \right)^{(p-2)/(p-1)} \\
&\qquad \cdot \left(\fint_{Q_j} |Du - Dw_j|^{p-1}\,dx\,dt \right)^{1/(p-1)} \\
&\leq c(\sigma) \bar{c}_3^{(1+\gamma)(p-2)} \bar{c}_1 \lambda \left[\frac{1}{\lambda} \frac{|\mu|(Q_{j-1})}{r_{j-1}^{N-1}} \right]^{[(1+\gamma)(p-2)+1](n+2)/[(p-1)n+p]}\,.
\end{aligned}$$

But now, as

$$\frac{[(1+\gamma)(p-2)+1]\,(n+2)}{(p-1)n+p} > \frac{(p-1)(n+2)}{(p-1)n+p} \geq 1 \tag{2.30}$$

precisely for $p \geq 2$, (2.24) implies

$$\fint_{Q_j} |D\bar{w}_j - D\bar{w}_{j-1}|^{(p-2)(1+\gamma)} |Du - Dw_j|\,dx\,dt \leq c \frac{|\mu|(Q_{j-1})}{r_{j-1}^{N-1}}$$

with $c \equiv c(n, p, \nu, \sigma)$. Therefore (2.29) gives us

$$\begin{aligned}
&\fint_{Q_j} |Du - Dw_j|\,dx\,dt \\
&\quad \leq c \fint_{Q_j} |D\bar{w}_j|^{(p-2)(1+\gamma)} |Du - Dw_j|\,dx\,dt + c \frac{|\mu|(Q_{j-1})}{r_{j-1}^{N-1}}\,.
\end{aligned} \tag{2.31}$$

We then continue to estimate the first term on the right in the above display. Applying Hölder's inequality, together with (2.21), and recalling that $\xi = 1 + 2\gamma$, we obtain for any $\alpha > 0$ that

$$
\begin{aligned}
&\fint_{Q_j} |D\bar{w}_j|^{(p-2)(1+\gamma)}|Du - Dw_j|\,dx\,dt \\
&\quad \le \fint_{Q_j} \left[\lambda^{2-p}\frac{(|Du| + |Dw_j|)^{p-2}|Du - Dw_j|^2}{(\alpha + |u - w_j|)^{\xi}}\right]^{1/2} \\
&\qquad\qquad \cdot \left[|D\bar{w}_j|^{(1+2\gamma)(p-2)}(\alpha + |u - w_j|)^{\xi}\right]^{1/2} dx\,dt \\
&\quad \le \lambda^{(2-p)/2}\left(\fint_{Q_j} \frac{(|Du| + |Dw_j|)^{p-2}|Du - Dw_j|^2}{(\alpha + |u - w_j|)^{\xi}}\,dx\,dt\right)^{1/2} \\
&\qquad\qquad \cdot \left(\fint_{Q_j} |D\bar{w}_j|^{\xi(p-2)}(\alpha + |u - w_j|)^{\xi}\,dx\,dt\right)^{1/2} \\
&\quad \le \left(\frac{\bar{c}_2}{\xi - 1}\right)^{1/2}\alpha^{(1-\xi)/2}\left[\frac{|\mu|(Q_j)}{r_j^N}\right]^{1/2} \\
&\qquad\qquad \cdot \left(\fint_{Q_j} |D\bar{w}_j|^{\xi(p-2)}(\alpha + |u - w_j|)^{\xi}\,dx\,dt\right)^{1/2}.
\end{aligned} \tag{2.32}
$$

As the choice of α is still in our disposal, we set

$$
\alpha := \left(\fint_{Q_j} |D\bar{w}_j|^{\xi(p-2)}|u - w_j|^{\xi}\,dx\,dt\right)^{1/\xi} + \delta \tag{2.33}
$$

for some small positive $\delta \in (0,1)$ to get

$$
\begin{aligned}
&\left(\fint_{Q_j} |D\bar{w}_j|^{\xi(p-2)}(\alpha + |u - w_j|)^{\xi}\,dx\,dt\right)^{1/2} \\
&\qquad \le 2\alpha^{\xi/2}\left(\fint_{Q_j} |D\bar{w}_j|^{\xi(p-2)}dx\,dt\right)^{1/2} + 2\alpha^{\xi/2}.
\end{aligned}
$$

Notice that since w_j belongs to the parabolic Sobolev space $L^p(T_j; W^{1,p}(B_j))$ and $\xi(p-1) < p$ by (2.27), we have that α is finite by Hölder's inequality. The presence in (2.33) of the parameter δ, which shall be sent to zero at the end of the proof, guarantees that α is positive. In the above display, the integral on the right can be estimated by mean of Lemma 2.3 as

$$
\begin{aligned}
&\fint_{Q_j} |D\bar{w}_j|^{\xi(p-2)}dx\,dt \\
&\quad \le c\fint_{Q_j} |D\bar{w}_{j-1} - D\bar{w}_j|^{\xi(p-2)}dx\,dt + c\fint_{Q_j} |D\bar{w}_{j-1}|^{\xi(p-2)}dx\,dt \\
&\quad \le c\bar{c}_3^{\xi(p-2)}\left[\frac{1}{\lambda}\frac{|\mu|(Q_{j-1})}{r_{j-1}^{N-1}}\right]^{\xi(p-2)(n+2)/[(p-1)n+p]} + cA^{\xi(p-2)} \le c\,,
\end{aligned}
$$

owing to (2.24) and (2.25), while the last constant c depends only on n, p, ν, σ, A. Thus (2.32), together with the last two displays, yields

$$\fint_{Q_j} |D\bar{w}_j|^{(p-2)(1+\gamma)}|Du - Dw_j| \,dx\,dt \le c\sqrt{\frac{\alpha}{r_j}}\left[\frac{|\mu|(Q_j)}{r_j^{N-1}}\right]^{1/2}$$

while applying Young's inequality together with an obvious estimation in turn gives

$$\fint_{Q_j} |D\bar{w}_j|^{(p-2)(1+\gamma)}|Du - Dw_j| \,dx\,dt \le \frac{c}{\beta}\frac{|\mu|(Q_{j-1})}{r_{j-1}^{N-1}} + \frac{\beta\alpha}{r_{j-1}}$$

for all $\beta \in (0,1)$, where $c \equiv c(n,p,\nu,\sigma,A)$. Inserting this into (2.31) leads to

$$\fint_{Q_j} |Du - Dw_j| \,dx\,dt \le \frac{c}{\beta}\frac{|\mu|(Q_{j-1})}{r_{j-1}^{N-1}} + \frac{\beta\alpha}{r_{j-1}} \tag{2.34}$$

again for all $\beta \in (0,1)$, where $c \equiv c(n,p,\nu,\sigma,A)$ is in particular independent of β. We then focus on α, which has been defined in (2.33), and split as follows:

$$\begin{aligned} \alpha &\le c\left(\fint_{Q_j} |D\bar{w}_{j-1} - D\bar{w}_j|^{\xi(p-2)}|u - w_j|^{\xi} \,dx\,dt\right)^{1/\xi} \\ &\quad + c\left(\fint_{Q_j} |D\bar{w}_{j-1}|^{\xi(p-2)}|u - w_j|^{\xi} \,dx\,dt\right)^{1/\xi} + \delta \quad =: \quad I_1 + I_2 + \delta\,. \end{aligned} \tag{2.35}$$

By (2.27), as $\xi(p-1) = \bar{q}$, we get by (2.22)-(2.23), together with Hölder's inequality, that

$$\begin{aligned} I_1 &\le c\left(\fint_{Q_j} |D\bar{w}_{j-1} - D\bar{w}_j|^{\bar{q}} \,dx\,dt\right)^{(p-2)/\bar{q}}\left(\fint_{Q_j} |u - w_j|^{\bar{q}} \,dx\,dt\right)^{1/\bar{q}} \\ &\le c\bar{c}_3^{p-1} r_{j-1}\lambda\left[\frac{1}{\lambda}\frac{|\mu|(Q_{j-1})}{r_{j-1}^{N-1}}\right]^{[(p-2)(n+2)+n+p]/[(p-1)n+p]}. \end{aligned}$$

Since

$$\frac{(p-2)(n+2)+n+p}{(p-1)n+p} = 1 + \frac{2(p-2)}{(p-1)n+p} \ge 1\,,$$

precisely for $p \ge 2$ we have, in view of (2.24), that

$$I_1 \le c r_{j-1}\left[\frac{|\mu|(Q_{j-1})}{r_{j-1}^{N-1}}\right]$$

for $c \equiv c(n,p,\nu,\sigma)$. On the other hand, using condition (2.25) we obtain

$$I_2 \le cA^{p-2}\left(\fint_{Q_j} |u - w_j|^{\xi} \,dx\,dt\right)^{1/\xi}.$$

Using Proposition 2.1, estimate (2.20), and Young's inequality we get

$$\begin{aligned} \left(\fint_{Q_j} |u - w_j|^{\xi} \,dx\,dt\right)^{1/\xi} &\le \left(\fint_{Q_j} |u - w_j|^{(n+1)/n} \,dx\,dt\right)^{n/(n+1)} \\ &\le c\left(\fint_{Q_j} |Du - Dw_j| \,dx\,dt\left[\sup_{t\in T_j}\fint_{B_j} |u - w_j| \,dx\right]^{1/n}\right)^{n/(n+1)} \\ &\le c\left(\fint_{Q_j} |Du - Dw_j| \,dx\,dt\right)^{n/(n+1)} r_{j-1}\left[\frac{|\mu|(Q_{j-1})}{r_{j-1}^{N-1}}\right]^{1/(n+1)} \end{aligned}$$

$$\leq c r_{j-1} \fint_{Q_j} |Du - Dw_j| \, dx \, dt + c r_{j-1} \left[\frac{|\mu|(Q_{j-1})}{r_{j-1}^{N-1}} \right].$$

Combining the estimates contained in the last three displays with (2.35) leads to

$$\alpha \leq c_* r_{j-1} \fint_{Q_j} |Du - Dw_j| \, dx \, dt + c_* r_{j-1} \left[\frac{|\mu|(Q_{j-1})}{r_{j-1}^{N-1}} \right] + \delta$$

with $c_* \equiv c_*(n, p, \nu, \sigma, A)$. Inserting this finally into (2.34) with $\beta = 1/(2c_*)$ and then reabsorbing terms and sending δ to zero finishes the proof. □

Next, a comparison estimate between w_j and v_j.

Lemma 2.5. *Let w_j and v_j be as in* (2.4) *and* (2.5)*, respectively, with $j \geq 0$. For every $A \geq 1$ there exists a constant $\bar{c}_5 \equiv \bar{c}_5(n, p, \nu, L, A)$ such that the following holds:*

$$\sup_{\frac{1}{2}Q_j} |Dw_j| + s \leq A\lambda$$

$$\Longrightarrow \quad \fint_{\frac{1}{2}Q_j} (|Dw_j| + |Dv_j|)^{p-2} |Dw_j - Dv_j|^2 \, dx \, dt$$

$$+ \fint_{\frac{1}{2}Q_j} |Dw_j - Dv_j|^p \, dx \, dt \leq \bar{c}_5 \, [\omega(r_j)]^2 \, \lambda^p . \tag{2.36}$$

Proof. The result is based on the following estimate:

$$\fint_{\frac{1}{2}Q_j} (|Dw_j| + |Dv_j|)^{p-2} |Dw_j - Dv_j|^2 \, dx \, dt \leq c \, [\omega(r_j)]^2 \fint_{\frac{1}{2}Q_j} (|Dw_j| + s)^p \, dx \, dt$$

that has been proved in [29, Lemma 4.3] and [1], with a constant $c \equiv c(n, p, \nu, L)$. The inequality in display (2.36) follows by trivially estimating

$$[\omega(r_j)]^2 \fint_{\frac{1}{2}Q_j} (|Dw_j| + s)^p \, dx \, dt \leq [\omega(r_j)]^2 \sup_{\frac{1}{2}Q_j} (|Dw_j| + s)^p \leq \bar{c}_5 \, [\omega(r_j)]^2 \, \lambda^p$$

and taking into account that $p \geq 2$. □

Finally, along the lines of Lemma 2.4, we yet prove another comparison estimate, this time between u and v_j.

Lemma 2.6. *Let u be as in Theorem 1.1 and let w_j and v_j be as in* (2.4) *and* (2.5)*, respectively, with $j \geq 1$. Suppose further that* (2.24) *holds together with*

$$\begin{cases} \displaystyle\sup_{\frac{1}{2}Q_j} |Dw_j| + s \leq A\lambda \\ \displaystyle\frac{\lambda}{A} \leq |Dw_{j-1}| \leq A\lambda \quad \text{in } Q_j \, , \end{cases} \tag{2.37}$$

for some $A \geq 1$. There exist positive constants $\bar{c}_6 \equiv \bar{c}_6(n, p, \nu, L, A)$ and $\bar{c}_7 \equiv \bar{c}_7(n, p, \nu, \sigma, A)$ such that the following inequality holds:

$$\fint_{\frac{1}{2}Q_j} |Du - Dv_j| \, dx \, dt \leq \bar{c}_6 \omega(r_j) \lambda + \bar{c}_7 \left[\frac{|\mu|(Q_{j-1})}{r_{j-1}^{N-1}} \right]. \tag{2.38}$$

Proof. We shall keep the notation introduced for Lemma 2.4; in particular, we recall (2.28). Inequality $(2.37)_1$ allows to use (2.36) so that

$$\fint_{\frac{1}{2}Q_j} (|Dw_j| + |Dv_j|)^{p-2} |Dw_j - Dv_j|^2 \, dx \, dt$$

$$+ \fint_{\frac{1}{2}Q_j} |Dw_j - Dv_j|^p \, dx \, dt \leq \bar{c}_5 \, [\omega(r_j)]^2 \, \lambda^p . \tag{2.39}$$

With $p' = p/(p-1)$, by $(2.37)_2$ we continue estimating as

$$
\begin{aligned}
\fint_{\frac{1}{2}Q_j} |Dw_j - Dv_j|\,dx\,dt &\leq A^{(p-2)/p'} \fint_{\frac{1}{2}Q_j} |D\bar{w}_{j-1}|^{(p-2)/p'} |Dw_j - Dv_j|\,dx\,dt \\
&\leq c \fint_{\frac{1}{2}Q_j} |D\bar{w}_j|^{(p-2)/p'} |Dw_j - Dv_j|\,dx\,dt \\
&\quad + c \fint_{\frac{1}{2}Q_j} |D\bar{w}_{j-1} - D\bar{w}_j|^{(p-2)/p'} |Dw_j - Dv_j|\,dx\,dt \qquad (2.40)
\end{aligned}
$$

with c depending only on p and A and we are using the notation in (2.28). As for the first term in the right hand side of (2.40), we notice that since $p \geq 2$ we have $(p-2)/2 \leq (p-2)/p'$ so that $(2.37)_1$ allows to estimate

$$
\begin{aligned}
&\fint_{\frac{1}{2}Q_j} |D\bar{w}_j|^{(p-2)/p'} |Dw_j - Dv_j|\,dx\,dt \\
&\qquad \leq A^{(p-2)^2/(2p)} \lambda^{(2-p)/2} \fint_{\frac{1}{2}Q_j} |Dw_j|^{(p-2)/2} |Dw_j - Dv_j|\,dx\,dt
\end{aligned}
$$

so that, using Hölder's inequality and (2.39) yields

$$
\begin{aligned}
&\fint_{\frac{1}{2}Q_j} |D\bar{w}_j|^{(p-2)/p'} |Dw_j - Dv_j|\,dx\,dt \\
&\qquad \leq c\lambda^{(2-p)/2} \left(\fint_{\frac{1}{2}Q_j} (|Dw_j| + |Dv_j|)^{p-2} |Dw_j - Dv_j|^2\,dx\,dt \right)^{1/2} \\
&\qquad \leq c\lambda^{(2-p)/2} \bar{c}_5^{1/2} \omega(r_j) \lambda^{p/2} = c\omega(r_j)\lambda \qquad (2.41)
\end{aligned}
$$

for $c \equiv c(n,p,\nu,L,A)$. The second term in the right hand side of (2.40) is estimated with the aid of Hölder's inequality, (2.23), (2.39) and finally Young's inequality, with conjugate exponents $(p/2, p/(p-2))$ when $p > 2$; this means

$$
\begin{aligned}
&\fint_{\frac{1}{2}Q_j} |D\bar{w}_{j-1} - D\bar{w}_j|^{(p-2)/p'} |Dw_j - Dv_j|\,dx\,dt \\
&\qquad \leq c \left(\fint_{Q_j} |D\bar{w}_{j-1} - D\bar{w}_j|^{p-2}\,dx\,dt \right)^{1/p'} \left(\fint_{\frac{1}{2}Q_j} |Dw_j - Dv_j|^p\,dx\,dt \right)^{1/p} \\
&\qquad \leq \bar{c}_3^{(p-2)/p'} \bar{c}_5^{1/p} \left[\frac{1}{\lambda} \frac{|\mu|(Q_{j-1})}{r_{j-1}^{N-1}} \right]^{[(p-2)/p](p-1)(n+2)/[n(p-1)+p]} [\omega(r_j)]^{2/p} \lambda \\
&\qquad \leq c \left[\frac{1}{\lambda} \frac{|\mu|(Q_{j-1})}{r_{j-1}^{N-1}} \right]^{(p-1)(n+2)/[n(p-1)+p]} \lambda + \omega(r_j)\lambda
\end{aligned}
$$

for $c \equiv c(n,p,\nu,L,\sigma,A)$. Thanks to (2.24) and (2.30), we get

$$
\fint_{\frac{1}{2}Q_j} |D\bar{w}_{j-1} - D\bar{w}_j|^{(p-2)/p'} |Dw_j - Dv_j|\,dx\,dt \leq \tilde{c} \left[\frac{|\mu|(Q_{j-1})}{r_{j-1}^{N-1}} \right] + \omega(r_j)\lambda
$$

with $\tilde{c} \equiv \tilde{c}(n,p,\nu,L,\sigma,A)$. By using the last inequality together with (2.40) and (2.41) we conclude with

$$
\fint_{\frac{1}{2}Q_j} |Dw_j - Dv_j|\,dx\,dt \leq c\omega(r_j)\lambda + \tilde{c} \left[\frac{|\mu|(Q_{j-1})}{r_{j-1}^{N-1}} \right]
$$

for $c \equiv c(n,p,\nu,L,A)$ and $\tilde{c} \equiv \tilde{c}(n,p,\nu,L,\sigma,A)$. Note next that the assumptions of this lemma fulfill also the assumptions of Lemma 2.4. Appealing then to (2.26) and triangle inequality finishes the proof. $\square$

2.5. **Comparison lemmas for SOLA.** In this section we show that the basic comparison Lemmas 2.2 and 2.3 hold for SOLA in a suitably modified way. This fact ultimately allows to prove Theorem 1.5 by mean of the same proofs already given for the other theorems when considered for energy solutions. Before starting, since the setting here is the one defined in Theorem 1.5, we assume the existence of functions u_h being local weak solutions to the equations considered in (1.22) and such that

$$Du_h \in L^p, \quad Du_h \to Du \ \text{ in } \ L^{p-1}, \quad u_h \to u \quad \text{and} \quad Du_h \to Du \ \ a.e. \tag{2.42}$$

hold. The measures μ_h weakly* converge to μ and (1.23) holds. Finally, since all the results we are interested in are local in nature, up to considering subsets compactly contained in Ω_T, we assume w.l.o.g. that the convergences in (2.42) are valid in the whole Ω_T.

Moreover, in this section we keep the notation already introduced in Section 2.2 about the cylinders Q_j, but the functions w_j will be defined in a different way: what matters here is not that they are solving Cauchy-Dirichlet problems as (2.4), but rather that they satisfy comparison estimates as in Lemmas 2.2 and 2.3. Indeed, once these lemmas are available for some suitably regular maps w_j then all the subsequent constructions can be replicated verbatim. More precisely, we have the following:

Lemma 2.7. *Let u be a SOLA to* (1.3) *as in Theorem 1.5 and let $\tilde\varepsilon \in (0, 1/(n+1)]$; moreover, let $\{Q_j\}$ be the sequence of cylinders considered in* (2.3). *Then there exists a sequence of functions w_j such that*

- $w_j \in C^0(T_j; L^2_{\rm loc}(B_j)) \cap L^p_{\rm loc}(T_j; W^{1,p}_{\rm loc}(B_j))$
- *w_j is a local weak solution to $\partial_t w_j - \operatorname{div} a(x,t,Dw_j) = 0$ in Q_j*
- *There exists a constant $\bar c_1 \equiv \bar c_1(n,p,\nu,\tilde\varepsilon)$ such that*

$$\left(\fint_{Q_j} |Du - Dw_j|^q \,dx\,dt\right)^{1/q} \le \bar c_1 \lambda \left[\frac{1}{\lambda}\frac{|\mu|(\lfloor Q_j\rfloor_{\rm par})}{r_j^{N-1}}\right]^{(n+2)/[(p-1)n+p]} \tag{2.43}$$

holds for any $0 < q \le p - 1 + 1/(n+1) - \tilde\varepsilon$

- *There exists a constant $\bar c_2 \equiv \bar c_2(n,p,\nu)$*

$$\fint_{Q_j} \frac{(|Du| + |Dw_j|)^{p-2}|Du - Dw_j|^2}{(\alpha + |u - w_j|)^{\xi}} \,dx\,dt \le \bar c_2 \frac{\alpha^{1-\xi}}{\xi - 1}\frac{|\mu|(\lfloor Q_j\rfloor_{\rm par})}{\lambda^{2-p} r_j^N} \tag{2.44}$$

holds for any $\alpha > 0$ and $\xi > 1$.

- *There exists a constant $\bar c_3 \equiv \bar c_3(n,p,\nu,\tilde\varepsilon,\sigma)$ such that the inequality*

$$\left(\fint_{Q_j} |u - w_j|^q \,dx\,dt\right)^{1/q} \le \bar c_3 r_{j-1} \lambda \left[\frac{1}{\lambda}\frac{|\mu|(\lfloor Q_{j-1}\rfloor_{\rm par})}{r_{j-1}^{N-1}}\right]^{(n+p)/[(p-1)n+p]} \tag{2.45}$$

holds whenever $0 < q \le p - 1 + p/n - \tilde\varepsilon$ and

$$\left(\fint_{Q_j} |Dw_{j-1} - Dw_j|^q \,dx\,dt\right)^{1/q} \le \bar c_3 \lambda \left[\frac{1}{\lambda}\frac{|\mu|(\lfloor Q_{j-1}\rfloor_{\rm par})}{r_{j-1}^{N-1}}\right]^{(n+2)/[(p-1)n+p]} \tag{2.46}$$

holds whenever $0 < q \le p - 1 + 1/(n+1) - \tilde\varepsilon$.

Proof. The proof goes via approximation; we prove (2.43)-(2.44), the proof of the remaining inequalities being similar. Let us fix a cylinder Q_j and consider the

sequence appearing in (2.42) and define

$$\tilde w_h \in C^0(T_j;L^2(B_j))\cap L^p(T_j;W^{1,p}(B_j))$$

be the unique solution to the Cauchy-Dirichlet problem

$$\begin{cases}\partial_t \tilde w_h - \operatorname{div} a(x,t,D\tilde w_h)=0 & \text{in } Q_j\\ \qquad\qquad\qquad\quad \tilde w_h = u_h & \text{on } \partial_{\mathrm{par}} Q_j\,.\end{cases}$$

Applying Lemma 2.2 to this context gives

$$\left(\fint_{Q_j}|Du_h-D\tilde w_h|^q\,dx\,dt\right)^{1/q}\le \bar c_1\lambda\left[\frac{1}{\lambda}\frac{|\mu_h|(Q_j)}{r_j^{N-1}}\right]^{(n+2)/[(p-1)n+p]} \tag{2.47}$$

for any $0<q\le p-1+1/(n+1)-\tilde\varepsilon$ and moreover

$$\fint_{Q_j}\frac{(|Du_h|+|D\tilde w_h|)^{p-2}|Du_h-D\tilde w_h|^2}{(\alpha+|u_h-\tilde w_h|)^{\xi}}\,dx\,dt\le \bar c_2\frac{\alpha^{1-\xi}}{\xi-1}\left[\frac{|\mu_h|(Q_j)}{\lambda^{2-p}r_j^N}\right]\,. \tag{2.48}$$

From (2.47) and (2.42) it follows that the sequence $\{D\tilde w_h\}$ is bounded in $L^{p-1}(Q_j)$. We now notice that $\tilde w_h$ is an energy solution and solves an equation with Dini-continuous coefficients (see Theorem 2.3 and subsequent Remark 2.1), moreover all these equations satisfy assumptions (1.4)-(1.5) uniformly in h. Hence, by interior regularity theory [18] and in particular by Corollary 2.1, it follows that the the maps $\{D\tilde w_h\}$ are locally uniformly bounded in L^∞. Again by interior regularity theory (see for instance the results in [29]) we have that the maps $\tilde w_h$ and $\{D\tilde w_h\}$ are locally uniformly equicontinuous in Q_j. Therefore, by Ascoli-Arzelà's theorem and a standard diagonal argument, we may assume that, up to a not relabeled sequence, there exists a limit map $w\in L^p_{\mathrm{loc}}(T_j;W^{1,p}_{\mathrm{loc}}(B_j))$ such that $\tilde w_h\to w$ in $L^p_{\mathrm{loc}}(T_j;W^{1,p}_{\mathrm{loc}}(B_j))$, $D\tilde w_h\to Dw$ and $\tilde w_h\to w$ locally uniformly and almost everywhere. As a consequence, w weakly solves $\partial_t w-\operatorname{div} a(x,t,Dw)=0$ in Q_j. At this stage the proof of (2.43) and (2.44) follows taking $w_j:=w$, letting $h\to\infty$ in (2.47)-(2.48) and using Fatou's lemma to deal with the left hand sides and (1.23) to deal with the right hand sides. Inequalities in displays (2.45)-(2.46) can be similarly proved by approximation starting by the analogs of (2.22)-(2.23), respectively, when written for u_h and $\tilde w_h$ as already done for (2.47)-(2.48). $\square$

3. Proof of Theorems 1.1 and 1.2

3.1. **Proof of Theorem 1.1.** The proof goes in several steps and involves a rather delicate induction argument. In the following we select a Lebesgue point of the spatial gradient $(x_0,t_0)\in\Omega_T$, i.e.,

$$\lim_{\varrho\to 0}\fint_{Q_\varrho(x_0,t_0)} Du\,dx\,dt = Du(x_0,t_0)\,. \tag{3.1}$$

Almost every point in Ω_T, with respect to the Lebesgue measure in $\mathbb{R}^{n+1}$, satisfies such a property (see [43, Chapter 1, Page 8]).

Step 1: Choice of constants and basic setup. With (x_0,t_0) being fixed at the beginning we shall verify (1.19) with $Q^\lambda_{2r}(x_0,t_0)$ instead of $Q^\lambda_r(x_0,t_0)$; this is of course causes no loss of generality. When finding constants c,R_0 such that (1.19) works, we choose positive numbers H_1,H_2 appearing in the lower bound for λ:
(3.2)

$$\lambda> H_1\left(\fint_{Q^\lambda_{2r}(x_0,t_0)}(|Du|+s)^{p-1}\,dx\,dt\right)^{1/(p-1)}+H_2\int_0^{2r}\frac{|\mu|(Q^\lambda_\varrho(x_0,t_0))}{\varrho^{N-1}}\frac{d\varrho}{\varrho}\,,$$

where $r \in (0, R_0/2)$ and R_0 is suitably small, to be determined in due course of the proof. The statement of Theorem 1.1 will be then proved with $c := 2\max\{H_1, H_2\}$. Without loss of generality, we may assume that λ is finite, since otherwise there is nothing to prove. In the rest of the proof certain constants will be *deliberately* chosen smaller/larger than necessary to emphasize the fact that their role is "to be very small/large". To begin with, having in mind to apply Theorems 2.3 and 2.4, we set

$$A := 1000^{4pN}\max\{c_2, c_3, 200\}\,, \quad B := 10^5\,, \quad \delta := 10^{-5}\,, \quad \bar{\varepsilon} := 4^{-(N+4)}\,, \tag{3.3}$$

where $c_2 \equiv c_2(n,p,\nu,L)$ and $c_3 \equiv c_3(n,p,\nu,L)$ are as in Corollary 2.1 and Theorem 2.2, respectively. In particular, A depends only on n,p,ν,L. Let $\sigma_1 \equiv \sigma_1(n,p,\nu,L,\omega(\cdot))$ and $\sigma_2 \equiv \sigma_2(n,p,\nu,L)$ be as in Theorems 2.3 and 2.4, respectively, both corresponding to the choices in (3.3). Set

$$\sigma := \min\{\sigma_1, \sigma_2, 1000^{-p/N}(p-1)^{-1/N}A^{-(p-2)/N}, 16^{-Np}\} \in (0, 1/4)\,. \tag{3.4}$$

The choices made above guarantee that all fixed parameters $A, B, \bar{\varepsilon}, \delta, \sigma$ depend only on $n,p,\nu,L,\omega(\cdot)$. Furthermore, let c_4 and α be as in Theorem 2.5, corresponding to the choice of A in (3.3). In this way they both depend only on n,p,ν,L. Next, let k be the smallest integer satisfying

$$c_4\sigma^{k\alpha} \le \frac{\sigma^N}{10^6} \quad \text{and} \quad k \ge 2\,. \tag{3.5}$$

As all of c_4, α and σ depend only on $n,p,\nu,L,\omega(\cdot)$, so does $k \equiv k(n,p,\nu,L,\omega(\cdot))$. We now proceed with the choice of H_1, H_2 and R_0. We set

$$H_1 := 100^{N/(p-1)}10^6\sigma^{-4N}\,. \tag{3.6}$$

Then, taking $\bar{c}_1 \equiv \bar{c}_1(n,p,\nu,1/(n+1)) \equiv \bar{c}_1(n,p,\nu)$ as in Lemma 2.2, $\bar{c}_3 \equiv \bar{c}_3(n,p,\nu,L,1/(n-1),\sigma) \equiv \bar{c}_3(n,p,\nu,L,\omega(\cdot))$ from Lemma 2.3, and $\bar{c}_7 \equiv \bar{c}_7(n,p,\nu,L,\omega(\cdot),\sigma) \equiv \bar{c}_7(n,p,\nu,L,\omega(\cdot))$ from Lemma 2.6 with A, σ as in (3.3)-(3.4), we set

$$H_2 := \left(2^N\sigma^{-N(k+5)}10^6 A\max\{\bar{c}_1, \bar{c}_3, \bar{c}_7\}\right)^{[n(p-1)+p]/(n+2)}\,. \tag{3.7}$$

Notice that since all the constants H_1 and H_2 are built on ultimately depend on $n,p,\nu,L,\omega(\cdot)$, we also have $H_1, H_2 \equiv H_1, H_2(n,p,\nu,L,\omega(\cdot))$. We then pass to the choice of R_0. Looking then at Lemmas 2.5 and 2.6 and taking $\bar{c}_5, \bar{c}_6$ corresponding to the choices of A and σ just made in (3.3)-(3.4), we take $R_2 \equiv R_2(n,p,\nu,L,\omega(\cdot))$ to be the largest positive number such that

$$\bar{c}_5^{1/p}\sigma^{-N(k+6)}[\omega(R_2)]^{2/p} + \bar{c}_6\sigma^{-(2N+1)}\int_0^{R_2}\omega(\varrho)\,\frac{d\varrho}{\varrho} \le \frac{1}{2^N 10^6} \tag{3.8}$$

is satisfied. Finally, with $R_1 \equiv R_1(n,p,\nu,L,\omega(\cdot))$ being as in Theorem 2.1 and Corollary 2.1, we let

$$R_0 := \min\{R_1, R_2\}/4\,. \tag{3.9}$$

From now on, with λ as in (3.2), $2r \le R_0$ and $Q^{\lambda}_{2r}(x_0,t_0) \subset \Omega_T$, we adopt and fix the basic set-up described at the beginning of Section 2.2. We therefore denote $r_j = \sigma^j r$ for $j \ge 0$, $Q_j \equiv Q^{\lambda}_{r_j}(x_0,t_0)$ and the comparison maps w_j and v_j defined in (2.4) and (2.5), respectively.

We finally record a few immediate consequences of the choices done so far for H_1, H_2, R_0. Observe that in (3.2) the latter potential term may be further estimated from below as follows (recall that $r_0 = r$):

$$\int_0^{2r}\frac{|\mu|(Q^{\lambda}_{\varrho}(x_0,t_0))}{\varrho^{N-1}}\,\frac{d\varrho}{\varrho}$$

$$
\begin{aligned}
&= \sum_{i=0}^{\infty} \int_{r_{i+1}}^{r_i} \frac{|\mu|(Q_\varrho^\lambda(x_0,t_0))}{\varrho^{N-1}}\,\frac{d\varrho}{\varrho} + \int_{r_0}^{2r_0} \frac{|\mu|(Q_\varrho^\lambda(x_0,t_0))}{\varrho^{N-1}}\,\frac{d\varrho}{\varrho} \\
&\geq \sum_{i=0}^{\infty} \frac{|\mu|(Q_{i+1})}{r_i^{N-1}} \int_{r_{i+1}}^{r_i} \frac{d\varrho}{\varrho} + \frac{|\mu|(Q_0)}{(2r_0)^{N-1}} \int_{r_0}^{2r_0} \frac{d\varrho}{\varrho} \\
&= \sigma^{N-1} \log\left(\frac{1}{\sigma}\right) \sum_{i=0}^{\infty} \frac{|\mu|(Q_{i+1})}{r_{i+1}^{N-1}} + \frac{\log 2}{2^{N-1}} \left[\frac{|\mu|(Q_0)}{r_0^{N-1}}\right] \geq \sigma^N \sum_{i=0}^{\infty} \frac{|\mu|(Q_i)}{r_i^{N-1}} \qquad (3.10)
\end{aligned}
$$

so that we conclude with

$$
\lambda > H_1 2^{-N/(p-1)} \left(\fint_{Q_0} (|Du|+s)^{p-1}\,dx\,dt\right)^{1/(p-1)} + H_2 \sigma^N \sum_{i=0}^{\infty} \frac{|\mu|(Q_i)}{r_i^{N-1}}\,. \qquad (3.11)
$$

The choice of H_1 in (3.6) guarantees for instance that

$$
s \leq \frac{\lambda}{600}\,. \qquad (3.12)
$$

Notice that in the decomposition (3.10) we can always assume that

$$
|\mu|(\partial_{\mathrm{par}} Q_i) = 0 \qquad (3.13)
$$

for every $i \geq 0$, since this amounts to change the integrand of the first integral in (3.10) in countably many points. The choice of H_2 in (3.7) together with (3.13), allows to deduce that

$$
\begin{aligned}
\sum_{i=0}^{\infty} \frac{|\mu|(Q_i)}{r_i^{N-1}} &= \sum_{i=0}^{\infty} \frac{|\mu|(\lfloor Q_i \rfloor_{\mathrm{par}})}{r_i^{N-1}} \\
&\leq \left(\frac{\sigma^{N(k+4)}}{2^N 10^6 A \max\{\bar c_1, \bar c_3, \bar c_7\}}\right)^{[n(p-1)+p]/(n+2)} \lambda \\
&\leq \frac{\sigma^{N(k+4)}}{2^N 10^6 A \max\{\bar c_1, \bar c_3, \bar c_7\}} \lambda \leq \lambda\,. \qquad (3.14)
\end{aligned}
$$

Notice that in the second estimate above we have used the fact that, since $p \geq 2$, then $n(p-1)+p \geq n+2$, and that the quantity in brackets is smaller than 1.

Recall now the choice of R_0 in (3.9) and observe that, if $r \equiv r_0 \in (0, R_2/4]$, then

$$
\begin{aligned}
\int_0^{R_2} \omega(\varrho)\,\frac{d\varrho}{\varrho} &= \sum_{i=0}^{\infty} \int_{r_{i+1}}^{r_i} \omega(\varrho)\,\frac{d\varrho}{\varrho} + \int_{r_0}^{R_2} \omega(\varrho)\,\frac{d\varrho}{\varrho} \\
&\geq \log\left(\frac{1}{\sigma}\right) \sum_{i=0}^{\infty} \omega(r_{i+1}) + \log 4\,\omega(r_0) \geq \sigma \sum_{i=0}^{\infty} \omega(r_i) \qquad (3.15)
\end{aligned}
$$

holds, so that (3.8) implies

$$
\bar c_5^{1/p} \sigma^{-N(k+4)} [\omega(r_0)]^{2/p} + \bar c_6 \sum_{i=0}^{\infty} \omega(r_i) \leq \frac{\sigma^{2N}}{2^N 10^6}\,. \qquad (3.16)
$$

Remark 3.1. When dealing with equations of the type in (1.20), i.e., equations without dependence on x, Theorem 2.3 is not any longer needed and therefore no dependence on $\omega(\cdot)$ occurs in the constants. Actually, we do not need the comparison functions w_j and we can just use v_j.

Remark 3.2. An ambiguity occurs when considering (3.13), since (3.13) is automatically satisfied when $\mu \in L^1$, which is precisely the case when proving Theorem 1.1. Here we nevertheless wanted to emphasize the fact that we can assume (3.13) at this stage also when μ is genuinely a measure. This fact will play a role when proving Theorem 1.5 below.

Step 2: Exit time argument. After having fixed the relevant constants in the previous step, we consider, for indexes $i \geq 1$, the quantities

$$C_i := \sum_{m=-1}^{0} \left(\fint_{Q_{i+m}} |Du|^{p-1} \, dx \, dt \right)^{1/(p-1)} + 2\sigma^{-N} \fint_{Q_i} |Du - (Du)_{Q_i}| \, dx \, dt \, .$$

By (3.6) and (3.11) it follows that

$$C_1 \leq 6\sigma^{-N-\frac{N}{p-1}} \left(\fint_{Q_0} |Du|^{p-1} \, dx \, dt \right)^{1/(p-1)} \leq \frac{\lambda}{1000} \, .$$

Furthermore, without loss of generality, we may assume that there is an exit time index $i_e \geq 1$ such that $C_i > \lambda/1000$ whenever $i > i_e$ and $C_{i_e} \leq \lambda/1000$. Indeed, if such an index does not exist, we have – in view of $C_1 \leq \lambda/1000$ – a subsequence $(i_j)_j$ of indexes such that $C_{i_j} \leq \lambda/1000$ as $j \to \infty$. But, as we assume that (x_0, t_0) is a Lebesgue point of Du, then also

$$|Du(x_0, t_0)| = \lim_{j \to \infty} |(Du)_{Q_{i_j}}| \leq \limsup_{j \to \infty} C_{i_j} \leq \frac{\lambda}{1000} \tag{3.17}$$

would hold and the proof would be complete. Thus, from now on, we shall work under the assumption

$$C_i > \frac{\lambda}{1000} \quad \text{for} \quad i \in \{i_e + 1, i_e + 2, \ldots\} \, , \quad \text{and} \quad C_{i_e} \leq \frac{\lambda}{1000} \, . \tag{3.18}$$

Note that in (3.17) we are using a limit computed on a sequence of shrinking intrinsic cylinders, therefore different from those considered in (3.1). On the other hand, define the set $\mathcal{L}_\lambda$ as

$$\mathcal{L}_\lambda := \left\{ (x_0, t_0) \in \Omega_T \ : \ \lim_{\varrho \to 0} \fint_{Q_\varrho^\lambda(x_0, t_0)} Du \, dx \, dt = Du(x_0, t_0) \right\}$$

for $\lambda > 0$. Basic properties of maximal operators - see for instance [43, Chapter 1, Page 8] - imply that this set is actually independent of λ and, in particular, $\mathcal{L}_\lambda = \mathcal{L}_1 =: \mathcal{L}$ for all $0 < \lambda < \infty$ and therefore (3.17) is completely justified in view of the assumed property in (3.1).

Step 3: Induction scheme. In order to prove (1.19) using (3.18), we apply an induction argument. To shorten the notation, we set, for $i \geq 0$,

$$E_i := \fint_{Q_i} |Du - (Du)_{Q_i}| \, dx \, dt \, , \qquad a_i := |(Du)_{Q_i}| \, . \tag{3.19}$$

We shall consider, in our iterative setting, the following conditions:

$$\mathrm{Ind}_1(j) : \qquad \left(\fint_{Q_{j-1}} |Du|^{p-1} \, dx \, dt \right)^{1/(p-1)} + \left(\fint_{Q_j} |Du|^{p-1} \, dx \, dt \right)^{1/(p-1)} \leq \lambda$$

for given integer $j \geq i_e$ and

$$\mathrm{Ind}_2(j) : \qquad \sum_{i=i_e+1}^{j} E_i \leq \frac{1}{2} \sum_{i=i_e}^{j-1} E_i + \frac{2\bar{c}_6}{\sigma^N} \sum_{i=i_e}^{j-1} \omega(r_i) \, \lambda + \frac{2\bar{c}_7}{\sigma^N} \sum_{i=i_e-1}^{j-2} \frac{|\mu|(Q_i)}{r_i^{N-1}}$$

for given integer $j > i_e$. The constants $\bar{c}_6, \bar{c}_7$ are those defined in Lemma 2.6 with the choices of A and σ made in (3.3) and (3.4), respectively; in this way, they ultimately depend on $n, p, \nu, L, \omega(\cdot)$.

The goal is now to show, by induction, that $\mathrm{Ind}_1(j)$ holds for every integer $j \geq i_e$ and $\mathrm{Ind}_2(j)$ holds for every integer $j > i_e$. Indeed, this will immediately prove Theorem 1.1 as

$$|Du(x_0,t_0)| = \lim_{j\to\infty} a_j \leq \limsup_{j\to\infty} \left(\fint_{Q_j} |Du|^{p-1}\,dx\,dt \right)^{1/(p-1)} \leq \lambda\,.$$

Now let us remark that $\mathrm{Ind}_1(i_e)$ is automatically satisfied since

$$\sum_{m=-1}^{0} \left(\fint_{Q_{i_e+m}} |Du|^{p-1}\,dx\,dt \right)^{1/(p-1)} \leq C_{i_e} \leq \frac{\lambda}{1000}\,. \tag{3.20}$$

Therefore, the rest of the proof develops according to the following scheme:

$$\mathrm{Ind}_1(i_e) \quad \Longrightarrow \quad \mathrm{Ind}_2(i_e+1)\,, \tag{3.21}$$

$$\begin{cases} \mathrm{Ind}_1(j) \\ \mathrm{Ind}_2(j) \end{cases} \quad \Longrightarrow \quad \mathrm{Ind}_2(j+1) \qquad \forall\ j > i_e \tag{3.22}$$

and

$$\begin{cases} \mathrm{Ind}_1(j) \\ \mathrm{Ind}_2(j+1) \end{cases} \quad \Longrightarrow \quad \mathrm{Ind}_1(j+1) \qquad \forall\ j \geq i_e\,. \tag{3.23}$$

We remark that in the following, unless otherwise stated, whenever we are considering $\mathrm{Ind}_1(j)$ we will do it in the general case $j \geq i_e$, while, when considering $\mathrm{Ind}_2(j)$ we will do it for all indexes $j > i_e$.

Step 4: Upper bounds implied by $\mathrm{Ind}_1(j)$. Towards the proof of the induction step, here we assume that $\mathrm{Ind}_1(j)$ holds for a certain index $j \geq i_e$ and exploit a few consequences of this. With the integer k being defined as in (3.5), observe that Lemma 2.2 and (3.14) imply that whenever $l \in \{0,1,\ldots,k+1\}$ the following holds:

$$\begin{aligned} &\left(\fint_{Q_{j-1+l}} |Du - Dw_{j-1}|^{p-1}\,dx\,dt \right)^{1/(p-1)} \\ &\qquad \leq \sigma^{-\frac{Nl}{p-1}} \left(\fint_{Q_{j-1}} |Du - Dw_{j-1}|^{p-1}\,dx\,dt \right)^{1/(p-1)} \\ &\qquad \leq \bar{c}_1 \sigma^{-N(k+1)} \lambda \left[\frac{1}{\lambda} \frac{|\mu|(Q_{j-1})}{r_{j-1}^{N-1}} \right]^{(n+2)/[(p-1)n+p]} \leq \frac{\sigma^N}{2^N 10^6}\lambda\,. \end{aligned} \tag{3.24}$$

Similarly, the inequalities

$$\left(\fint_{Q_{j+l}} |Du - Dw_j|^{p-1}\,dx\,dt \right)^{1/(p-1)} \leq \frac{\sigma^N}{2^N 10^6}\lambda \tag{3.25}$$

hold as well for $l \in \{0,1,\ldots,k+1\}$. Using (3.24) with $l = 0$ and $\mathrm{Ind}_1(j)$ we get

$$\begin{aligned} &\left(\fint_{Q_{j-1}} |Dw_{j-1}|^{p-1}\,dx\,dt \right)^{1/(p-1)} \\ &\qquad \leq \frac{\sigma^N}{2^N 10^6}\lambda + \left(\fint_{Q_{j-1}} |Du|^{p-1}\,dx\,dt \right)^{1/(p-1)} \leq 2\lambda\,. \end{aligned}$$

By Corollary 2.1 (applied to w_{j-1}) we then obtain, also recalling (3.12), that

$$\sup_{\frac{1}{2}Q_{j-1}} |Dw_{j-1}|+s \leq c_2(\lambda+s)+c_2\lambda^{2-p} \fint_{Q_{j-1}} (|Dw_{j-1}|+s)^{p-1}\,dx\,dt \leq A\lambda\,. \tag{3.26}$$

In a completely similar way, by using this time (3.25) instead of (3.24), and again $\mathrm{Ind}_1(j)$, we can prove

$$\left(\fint_{Q_j} |Dw_j|^{p-1}\,dx\,dt\right)^{1/(p-1)} \leq 2\lambda \tag{3.27}$$

and then

$$\sup_{\frac{1}{2}Q_j} |Dw_j| + s \leq A\lambda\,. \tag{3.28}$$

Furthermore, keeping (3.16) in mind, estimate (3.28) and Lemma 2.5 allow us to deduce

$$\left(\fint_{\frac{1}{2}Q_j} |Dw_j - Dv_j|^{p-1}\,dx\,dt\right)^{1/(p-1)} \leq \tilde{c}_5^{1/p}\,[\omega(r_j)]^{2/p}\,\lambda \leq \frac{\sigma^{N(k+6)}}{2^N 10^6}\lambda \tag{3.29}$$

so that, for $l \in \{1,\ldots,k+1\}$ it holds that

$$\left(\fint_{Q_{j+l}} |Dw_j - Dv_j|^{p-1}\,dx\,dt\right)^{1/(p-1)} \leq \frac{\sigma^{2N}}{2^N 10^6}\lambda\,.$$

Combining the above estimate with (3.25) gives

$$\left(\fint_{Q_{j+l}} |Du - Dv_j|^{p-1}\,dx\,dt\right)^{1/(p-1)} \leq \frac{\sigma^N}{10^6}\lambda \tag{3.30}$$

for all $l \in \{1,\ldots,k+1\}$. Next, observe that (3.27) and (3.29) imply

$$\begin{aligned}
&\left(\fint_{\frac{1}{2}Q_j} |Dv_j|^{p-1}\,dx\,dt\right)^{1/(p-1)}\\
&\leq \left(\fint_{\frac{1}{2}Q_j} |Dw_j|^{p-1}\,dx\,dt\right)^{1/(p-1)} + \left(\fint_{\frac{1}{2}Q_j} |Dw_j - Dv_j|^{p-1}\,dx\,dt\right)^{1/(p-1)}\\
&\leq 2^{N+1}\lambda + \lambda \leq 4^N\lambda
\end{aligned}$$

so that Theorem 2.2, (3.3) and (3.12) finally yield

$$\sup_{\frac{1}{4}Q_j} |Dv_j| + s \leq c_3(\lambda+s) + c_3\lambda^{2-p}\fint_{\frac{1}{2}Q_j} (|Dv_j|+s)^{p-1}\,dx\,dt \leq A\lambda\,. \tag{3.31}$$

Step 5: Lower bounds implied by $\mathrm{Ind}_1(j)$. Here we still exploit a few consequences of assuming $\mathrm{Ind}_1(j)$ for some $j \geq i_e$. In particular, we derive suitable lower bounds for Dw_{j-1} and Dv_j. First, let us first show a few oscillation reduction estimates. Since we have already established the upper bound for Dw_{j-1} in (3.26), Theorem 2.3 applied to w_{j-1} (with $\delta = 10^{-5}$ as in (3.3) and recalling also the choice made in (3.4) that guarantees $\sigma \leq \sigma_1$ so that $Q_j \subset \sigma_1 Q_{j-1}$) gives

$$\operatorname*{osc}_{Q_j} Dw_{j-1} \leq \frac{\lambda}{10^5}\,. \tag{3.32}$$

For Dv_j we instead have (3.31) and hence Theorem 2.5 and (3.5) imply

$$\operatorname*{osc}_{Q_{j+k}} Dv_j \leq c_4\sigma^{\alpha k}\lambda \leq \frac{\sigma^N}{10^6}\lambda\,.$$

Using this together with (2.2) and (3.30) gives

$$
\begin{aligned}
&2\sigma^{-N} \fint_{Q_{j+k}} |Du-(Du)_{Q_{j+k}}|\,dx\,dt\\
&\quad\le 4\sigma^{-N} \fint_{Q_{j+k}} |Dv_j-(Dv_j)_{Q_{j+k}}|\,dx\,dt + 4\sigma^{-N}\fint_{Q_{j+k}} |Du-Dv_j|\,dx\,dt\\
&\quad\le 4\sigma^{-N} \operatorname*{osc}_{Q_{j+k}} Dv_j + 4\sigma^{-N}\fint_{Q_{j+k}} |Du-Dv_j|\,dx\,dt \le \frac{\lambda}{10^5}\,.
\end{aligned}
$$

But, as $C_{j+k} > \lambda/1000$ for $j \ge i_e$, we then necessarily have that

$$
\sum_{m=-1}^{0} \left(\fint_{Q_{j+m+k}} |Du|^{p-1}\,dx\,dt \right)^{1/(p-1)} \ge \frac{\lambda}{2000}\,. \tag{3.33}
$$

Next, again by using triangle inequality and (3.24) (for $l = k, k+1$), we have

$$
\begin{aligned}
&\sum_{m=-1}^{0} \left(\fint_{Q_{j+m+k}} |Du|^{p-1}\,dx\,dt \right)^{1/(p-1)}\\
&\qquad \le \frac{\lambda}{10^6} + \sum_{m=-1}^{0} \left(\fint_{Q_{j+m+k}} |Dw_{j-1}|^{p-1}\,dx\,dt \right)^{1/(p-1)}
\end{aligned} \tag{3.34}
$$

so that, as $k \ge 2$ and (3.33) holds, we also get

$$
\begin{aligned}
2\sup_{Q_j} |Dw_{j-1}| \;&\ge\; \sum_{m=-1}^{0} \left(\fint_{Q_{j+m+k}} |Dw_{j-1}|^{p-1}\,dx\,dt \right)^{1/(p-1)}\\
&\ge\; \frac{\lambda}{2000} - \frac{\lambda}{10^6} \ge \frac{3\lambda}{10^4}\,.
\end{aligned} \tag{3.35}
$$

Arguing as for (3.34), and using (3.30), we this time have

$$
\begin{aligned}
&\sum_{m=-1}^{0} \left(\fint_{Q_{j+m+k}} |Du|^{p-1}\,dx\,dt \right)^{1/(p-1)}\\
&\qquad \le \frac{\lambda}{10^6} + \sum_{m=-1}^{0} \left(\fint_{Q_{j+m+k}} |Dv_j|^{p-1}\,dx\,dt \right)^{1/(p-1)}
\end{aligned}
$$

so that, using again (3.33), and recalling that $k \ge 2$, we conclude with

$$
2\sup_{Q_{j+1}} |Dv_j| \ge \sum_{m=-1}^{0} \left(\fint_{Q_{j+m+k}} |Dv_j|^{p-1}\,dx\,dt \right)^{1/(p-1)} \ge \frac{3\lambda}{10^4}\,. \tag{3.36}
$$

The inequality in display (3.35) yields the existence of a point $(\tilde{x},\tilde{t}) \in Q_j$ such that $|Dw_{j-1}(\tilde{x},\tilde{t})| \ge \lambda/10^4$ and therefore the oscillation control in (3.32), recalling also (3.26) and that $Q_j \subset \frac{1}{2}Q_{j-1}$, gives

$$
\frac{\lambda}{A} \le \frac{\lambda}{10^5} \le |Dw_{j-1}| \le A\lambda \qquad \text{in } Q_j\,. \tag{3.37}
$$

Finally, (3.36) and (3.31), and the fact that $Q_{j+1} \subset \frac{1}{4}Q_j$, give

$$
\frac{\lambda}{B} \equiv \frac{\lambda}{10^5} \le \sup_{Q_{j+1}} |Dv_j| \le \sup_{\frac{1}{4}Q_j} |Dv_j| + s \le A\lambda\,. \tag{3.38}
$$

Step 6: Further consequences of $\mathrm{Ind}_1(j)$. We again assume $\mathrm{Ind}_1(j)$ for an arbitrary chosen $j \geq i_e$. Now, on one hand (3.14), (3.28) and (3.37) allow to apply Lemma 2.6 and obtain

$$\fint_{\frac{1}{2}Q_j} |Du - Dv_j| \, dx \, dt \leq \bar{c}_6 \omega(r_j)\lambda + \bar{c}_7 \left[\frac{|\mu|(Q_{j-1})}{r_{j-1}^{N-1}} \right] . \tag{3.39}$$

On the other hand, recalling that $\sigma Q_j = Q_{j+1}$, Theorem 2.4 with $\bar{\varepsilon} = 4^{-(N+4)}$ is at our disposal by (3.38) and it yields

$$\fint_{Q_{j+1}} |Dv_j - (Dv_j)_{Q_{j+1}}| \, dx \, dt \leq 4^{-(N+4)} \fint_{\frac{1}{4}Q_j} |Dv_j - (Dv_j)_{\frac{1}{4}Q_j}| \, dx \, dt \, . \tag{3.40}$$

Now, by (3.39), recalling the definitions in (3.19) and using (2.2) repeatedly, both

$$\begin{aligned}
&\fint_{\frac{1}{4}Q_j} |Dv_j - (Dv_j)_{\frac{1}{4}Q_j}| \, dx \, dt \\
&\qquad \leq 2^{2N+1} \fint_{Q_j} |Du - (Du)_{Q_j}| \, dx \, dt + 2^{N+2} \fint_{\frac{1}{2}Q_j} |Du - Dv_j| \, dx \, dt \\
&\qquad \leq 2^{2N+1} E_j + 2^{N+2} \bar{c}_6 \omega(r_j)\lambda + 2^{N+2} \bar{c}_7 \left[\frac{|\mu|(Q_{j-1})}{r_{j-1}^{N-1}} \right]
\end{aligned}$$

and

$$\begin{aligned}
&2 \fint_{Q_{j+1}} |Dv_j - (Dv_j)_{Q_{j+1}}| \, dx \, dt \\
&\qquad \geq \fint_{Q_{j+1}} |Du - (Du)_{Q_{j+1}}| \, dx \, dt - \sigma^{-N} \fint_{\frac{1}{2}Q_j} |Du - Dv_j| \, dx \, dt \\
&\qquad \geq E_{j+1} - \sigma^{-N} \bar{c}_6 \omega(r_j)\lambda - \sigma^{-N} \bar{c}_7 \left[\frac{|\mu|(Q_{j-1})}{r_{j-1}^{N-1}} \right]
\end{aligned}$$

hold. Combining the inequalities in the last three displays gives

$$E_{j+1} \leq \frac{1}{2} E_j + \frac{2\bar{c}_6}{\sigma^N} \omega(r_j)\lambda + \frac{2\bar{c}_7}{\sigma^N} \left[\frac{|\mu|(Q_{j-1})}{r_{j-1}^{N-1}} \right] . \tag{3.41}$$

Step 7: Verification of $\mathrm{Ind}_2(i_e+1)$ and $\mathrm{Ind}_2(j+1)$. Here we prove (3.21) and (3.22). The outcome of Step 6 is that (3.41) holds whenever $\mathrm{Ind}_1(j)$ holds, for every $j \geq i_e$, therefore, since $\mathrm{Ind}_1(i_e)$ holds as established in (3.20), taking $j = i_e$ in (3.41) yields $\mathrm{Ind}_2(i_e+1)$. We now come to the proof of (3.22), that is the validity of $\mathrm{Ind}_2(j+1)$; this simply follows by summing (3.41), that holds by the assumed validity of $\mathrm{Ind}_1(j)$, to the inequality yielded by the definition of $\mathrm{Ind}_2(j)$, which is also assumed in (3.22).

Step 8: Bounds for a_j and E_j. It remains to prove (3.23). For this we now assume that $\mathrm{Ind}_1(j)$ and $\mathrm{Ind}_2(j+1)$ hold for some $j \geq i_e$ and derive bounds for the quantities defined in (3.19). By $\mathrm{Ind}_2(j+1)$ and easy manipulations, we note that (3.14), (3.16), and $C_{i_e} \leq \lambda/1000$ imply

$$\begin{aligned}
\sum_{i=i_e}^{j+1} E_i &\leq 2E_{i_e} + \frac{4\bar{c}_6}{\sigma^N} \sum_{i=0}^{\infty} \omega(r_i)\, \lambda + \frac{4\bar{c}_7}{\sigma^N} \sum_{i=0}^{\infty} \frac{|\mu|(Q_i)}{r_i^{N-1}} \\
&\leq 2E_{i_e} + \frac{\sigma^N \lambda}{1000} \leq \sigma^N C_{i_e} + \frac{\sigma^N \lambda}{1000} \leq \frac{\sigma^N \lambda}{500} \, .
\end{aligned} \tag{3.42}$$

Using this together with the obvious estimation (valid whenever $i \geq 0$)

$$
\begin{aligned}
a_{i+1} - a_i &\leq \fint_{Q_{i+1}} |Du - (Du)_{Q_i}| \, dx \, dt \\
&\leq \frac{|Q_i|}{|Q_{i+1}|} \fint_{Q_i} |Du - (Du)_{Q_i}| \, dx \, dt = \sigma^{-N} E_i
\end{aligned}
$$

we get, after telescoping the previous inequalities and using (3.42), that

$$
a_{l+1} \leq a_{i_e} + \sigma^{-N} \sum_{i=i_e}^{l} E_i \leq \frac{\lambda}{1000} + \frac{\lambda}{500} \leq \frac{\lambda}{200} \tag{3.43}
$$

for all $l \in \{i_e, \ldots, j\}$. Here we also used that, thanks to (3.18), we have

$$
a_{i_e} \leq \left(\fint_{Q_{i_e}} |Du|^{p-1} \, dx \, dt \right)^{1/(p-1)} \leq C_{i_e} \leq \frac{\lambda}{1000} \, . \tag{3.44}
$$

Step 9: Verification of $\mathbf{Ind}_1(j+1)$. Here we prove (3.23), thereby concluding the proof. We actually have to prove that

$$
\sum_{m=-1}^{0} \left(\fint_{Q_{j+m+1}} |Du|^{p-1} \, dx \, dt \right)^{1/(p-1)} \leq \lambda \, . \tag{3.45}
$$

To this end, we estimate using (3.24) (with $l = 1, 2$) as follows:

$$
\begin{aligned}
&\sum_{m=-1}^{0} \left(\fint_{Q_{j+m+1}} |Du|^{p-1} \, dx \, dt \right)^{1/(p-1)} \\
&\quad \leq \sum_{m=-1}^{0} \left(\fint_{Q_{j+m+1}} |Du - Dw_{j-1}|^{p-1} \, dx \, dt \right)^{1/(p-1)} \\
&\quad\quad + \sum_{m=-1}^{0} \left(\fint_{Q_{j+m+1}} |Dw_{j-1}|^{p-1} \, dx \, dt \right)^{1/(p-1)} \\
&\quad \leq \frac{\sigma^N}{2^N 10^5} \lambda + \sum_{m=-1}^{0} \left(\fint_{Q_{j+m+1}} |Dw_{j-1}|^{p-1} \, dx \, dt \right)^{1/(p-1)} \, .
\end{aligned} \tag{3.46}
$$

We further estimate the latter term on the right hand side by simply applying triangle inequality as follows:

$$
\begin{aligned}
&\sum_{m=-1}^{0} \fint_{Q_{j+m+1}} |Dw_{j-1}|^{p-1} \, dx \, dt \\
&\leq \sum_{m=-1}^{0} \fint_{Q_{j+m+1}} |Dw_{j-1}|^{p-2} |(Du)_{Q_{j+m+1}}| \, dx \, dt \\
&\quad + \sum_{m=-1}^{0} \fint_{Q_{j+m+1}} |Dw_{j-1}|^{p-2} \left(|Du - (Du)_{Q_{j+m+1}}| + |Du - Dw_{j-1}| \right) \, dx \, dt \, .
\end{aligned}
$$

Using Young's inequality with conjugate exponents $(p-1)/(p-2)$ and $p-1$ (only when $p > 2$) we get

$$
\sum_{m=-1}^{0} \fint_{Q_{j+m+1}} |Dw_{j-1}|^{p-2} |(Du)_{Q_{j+m+1}}| \, dx \, dt
$$

$$\leq \frac{p-2}{p-1} \sum_{m=-1}^{0} \fint_{Q_{j+m+1}} |Dw_{j-1}|^{p-1}\, dx\, dt$$
$$+ \frac{1}{p-1} \sum_{m=-1}^{0} |(Du)_{Q_{j+m+1}}|^{p-1} \,.$$

Matching the inequalities in the last two displays, reabsorbing terms, and using (3.26), yields

$$\sum_{m=-1}^{0} \fint_{Q_{j+m+1}} |Dw_{j-1}|^{p-1}\, dx\, dt \leq \sum_{m=-1}^{0} |(Du)_{Q_{j+m+1}}|^{p-1}$$
$$+(p-1)(A\lambda)^{p-2} \sum_{m=-1}^{0} \fint_{Q_{j+m+1}} \big(|Du-(Du)_{Q_{j+m+1}}| + |Du - Dw_{j-1}|\big)\, dx\, dt \,.$$

Notice that we have used that $Q_j \subset \frac{1}{2}Q_{j-1}$ in order to apply (3.26). Now we estimate all the terms in the right hand side of the above inequality. Using (3.43) and (3.44) (this last one only when $j = i_e$) from Step 7 we have

$$\sum_{m=-1}^{0} |(Du)_{Q_{j+m+1}}|^{p-1} = a_{j+1}^{p-1} + a_j^{p-1} \leq 2\left(\frac{\lambda}{200}\right)^{p-1} \,,$$

while using (3.42) gives

$$\sum_{m=-1}^{0} \fint_{Q_{j+m+1}} |Du-(Du)_{Q_{j+m+1}}|\, dx\, dt = E_{j+1} + E_j \leq \frac{\sigma^N \lambda}{500} \,.$$

Finally, using (3.24) (with $l = 1, 2$) and Hölder's inequality yields

$$\sum_{m=-1}^{0} \fint_{Q_{j+m+1}} |Du - Dw_{j-1}|\, dx\, dt \leq \frac{\sigma^N}{2^N 10^5}\lambda \,.$$

Connecting the inequalities in the last four displays, and recalling the very definitions of A and σ in (3.3) and (3.4) respectively, gives us

$$\sum_{m=-1}^{0} \fint_{Q_{j+m+1}} |Dw_{j-1}|^{p-1}\, dx\, dt \leq \left(\frac{\lambda}{8}\right)^{p-1} \,.$$

Inserting the last inequality into (3.46) leads to (3.45). Therefore (3.23) is verified and the proof of Theorem 1.1 is complete.

3.2. **Proof of Theorem 1.2.** With the standard cylinder $Q_r \equiv Q_r(x_0, t_0) \subset \Omega_T$ being fixed in the statement, let us consider the function $h(\lambda) := \lambda - cA(\lambda)$ where

$$\begin{aligned} A(\lambda) \quad &:= \quad \lambda^{\frac{p-2}{p-1}} \left(\frac{1}{|Q_r|}\int_{Q_r^\lambda} (|Du|+s+1)^{p-1}\, dx\, dt\right)^{1/(p-1)} + \mathbf{I}^{\mu}_{1,\lambda}(x_0,t_0;r) \\ &:= \quad \left(\fint_{Q_r^\lambda} (|Du|+s+1)^{p-1}\, dx\, dt\right)^{1/(p-1)} + \mathbf{I}^{\mu}_{1,\lambda}(x_0,t_0;r) \end{aligned}$$

and $c > 1$ is again the constant appearing in Theorem 1.1; it depends only on $n, p, \nu, L, \omega(\cdot)$. Here it is $Q_r^\lambda \equiv Q_r^\lambda(x_0, t_0)$. We are actually considering the function $h(\cdot)$ to be defined for all those positive λ which are such that $Q_r^\lambda \subset \Omega_T$; observe that

since $p \geq 2$ then the domain of definition of $h(\cdot)$ includes $[1,\infty)$ as $Q_r^\lambda \subset Q_r \subset \Omega_T$ when $\lambda \geq 1$. Moreover, observe that again when $\lambda \geq 1$ we have

$$A(\lambda) \leq \lambda^{\frac{p-2}{p-1}} \left(\fint_{Q_r} (|Du| + s + 1)^{p-1} \, dx \, dt \right)^{1/(p-1)} + \mathbf{I}_1^\mu(x_0, t_0; r) \,. \tag{3.47}$$

The function $h(\cdot)$ is obviously continuous and moreover $h(1) < 0$ since $c > 1$ and $A(\lambda) \geq 1$. On the other hand, (3.47) implies that $h(\lambda) \to \infty$ as $\lambda \to \infty$. It follows that there exists a number $\lambda > 1$ such that $h(\lambda) = 0$ and therefore $\lambda = cA(\lambda)$. In particular, λ satisfies (1.19). Therefore we can apply Theorem 1.1 that gives

$$\lambda + |Du(x_0, t_0)| \leq 2\lambda = 2cA(\lambda) \,.$$

Using in turn Young's inequality with conjugate exponents $(p-1, (p-1)/(p-2))$ when $p > 2$ and (3.47), we have

$$2cA(\lambda) \leq \frac{\lambda}{2} + \tilde{c} \fint_{Q_r} (|Du| + s + 1)^{p-1} \, dx \, dt + 2c\mathbf{I}_1^\mu(x_0, t_0; r)$$

where $\tilde{c}$ depends only on $n, p, \nu, L, \omega(\cdot)$. The proof follows connecting the inequalities in the last two displays.

4. Proof of Theorems 1.3 and 1.4

4.1. **Proof of Theorem 1.4.** The proof consists of several steps; some of the arguments of the proof of Theorem 1.1 will be re-proposed in order to, this time, control the degeneracy rate of the equation.

Step 1: Basic setup and smallness conditions. Since we are assuming that $\mathbf{I}_1^\mu(x_0, t_0; r)$ is locally bounded for some $r > 0$ then by Theorem 1.2 we have that Du is locally bounded in Ω_T, too. Moreover, since we are proving a local statement, up to passing to open subsets compactly contained in Ω_T, we can assume w.l.o.g. that the gradient is globally bounded, therefore letting

$$\lambda := \|Du\|_{L^\infty(\Omega_T)} + s + 1 < \infty \,. \tag{4.1}$$

From now on our analysis will proceed on cylinders of the type $Q_r^\lambda \subset \Omega_T$. We shall prove that for every $\varepsilon \in (0,1)$, there exists a radius $r_\varepsilon \equiv r_\varepsilon(n, p, \nu, L, \omega(\cdot), \mu(\cdot), \varepsilon) > 0$ such that

$$E(Du, Q_\varrho^\lambda) = \fint_{Q_\varrho^\lambda} |Du - (Du)_{Q_\varrho^\lambda}| \, dx \, dt < \lambda\varepsilon \tag{4.2}$$

holds whenever $\varrho \in (0, r_\varepsilon]$ and $Q_\varrho^\lambda \subset \Omega_T$. Once this fact is proved the VMO-regularity of Du in the sense of Theorem 1.4 follows by an easy change-of-variables argument as λ is now fixed in (4.1).

Towards the application of Theorems 2.3-2.4, and with $c_2 \equiv c_2(n, p, \nu, L)$ and $c_3 \equiv c_3(n, p, \nu, L)$ being as in Corollary 2.1 and Theorem 2.2, respectively, we start fixing

$$A := \frac{1000^{4pN} \max\{c_2, c_3, 200\}}{\varepsilon} \,, \quad B := \frac{10^5}{\varepsilon} \,, \quad \delta := \frac{\varepsilon}{10^5} \,, \quad \bar{\varepsilon} := \frac{\varepsilon}{10^{5N}} \,. \tag{4.3}$$

With the choice in (4.3) we determine $\sigma_1 \equiv \sigma_1(n, p, \nu, L, \omega(\cdot), \varepsilon)$ and $\sigma_2 \equiv \sigma_2(n, p, \nu, L, \varepsilon)$ from Theorems 2.3 and 2.4, respectively. Next, we fix

$$\sigma := \min\{\sigma_1, \sigma_2, 16^{-N}\} \in (0, 1/4) \tag{4.4}$$

and notice that σ only depends on $n, p, \nu, L, \omega(\cdot), \varepsilon$. Such choices, looking at Lemmas 2.2, 2.5 and 2.6, determine the constants $\bar{c}_1, \bar{c}_5, \bar{c}_6$ and $\bar{c}_7$ as depending only on $n, p, \nu, L, \omega(\cdot), \varepsilon$. We again need some limitations on the size of the radii considered; here $R_1 \equiv R_1(n, p, \nu, L, \omega(\cdot))$ still denotes the radius considered in Theorem 2.1 and

Corollary 2.1. Moreover, we select a new radius $R_3 \equiv R_3(n,p,\nu,L,\omega(\cdot),\mu(\cdot),\varepsilon)$ in such a way that the following *smallness conditions* hold:

$$\sup_{0<\varrho\leq R_3}\ \sup_{(x,t)\in\Omega_T}\left[\frac{|\mu|(Q^\lambda_\varrho(x,t))}{\lambda\varrho^{N-1}}\right]^{(n+2)/[(p-1)n+p]} \leq \frac{\sigma^{3N}\varepsilon}{10^6\bar{c}_1\bar{c}_5\bar{c}_6\bar{c}_7} \tag{4.5}$$

and

$$[\omega(R_3)]^{2/p} \leq \frac{\sigma^{3N}\varepsilon}{10^6\bar{c}_1\bar{c}_5\bar{c}_6\bar{c}_7} \tag{4.6}$$

and we this time set $\bar{R}_0 := \min\{R_1,R_3\}/4$; as a consequence we have $\bar{R}_0 \equiv \bar{R}_0(n,p,\nu,L,\omega(\cdot),\mu(\cdot),\varepsilon)$. Notice only that the possibility of gaining (4.5) follows directly by the assumption (1.21), and from the fact that, since $\lambda\geq 1$ and $p\geq 2$, then $Q^\lambda_\varrho(x,t)\subset Q_\varrho(x,t)$. We now fix a cylinder $Q\equiv Q^\lambda_r(x_0,t_0)\subset\Omega_T$ with $r\in(\sigma\bar{R}_0,\bar{R}_0]$, and accordingly to the setup defined in Section 2.2 define the chain of shrinking parabolic intrinsic cylinders as follows:

$$Q_j \equiv Q^\lambda_{r_j}(x_0,t_0)\,,\qquad r_j=\sigma^j r,\qquad \text{where } \ r\in(\sigma\bar{R}_0,\bar{R}_0]\,, \tag{4.7}$$

for every integer $j\geq 0$. The related comparison solutions w_j and v_j are accordingly defined as in (2.4) and (2.5), respectively; finally, we denote

$$E_j := \fint_{Q_j}|Du-(Du)_{Q_j}|\,dx\,dt\,.$$

We shall preliminary prove that

$$E_{j+1}<\lambda\varepsilon \qquad\qquad \forall\, j\in\mathbb{N}\cap[1,\infty)\,. \tag{4.8}$$

Now, it is obvious that if for a given $j\geq 1$ it occurs that

$$\left(\fint_{Q_{j+1}}|Du|^{p-1}\,dx\,dt\right)^{1/(p-1)} < \frac{\lambda\varepsilon}{50}\,, \tag{4.9}$$

then (4.8) holds, therefore we can confine ourselves to assume that (4.9) does not hold. To prove (4.8) in this last case we start proving the following implication, which is valid whenever $j\geq 1$:

$$\begin{aligned}&\left(\fint_{Q_{j+1}}|Du|^{p-1}\,dx\,dt\right)^{1/(p-1)} \geq \frac{\lambda\varepsilon}{50}\\ &\qquad\Longrightarrow E_{j+1}\leq\frac{\varepsilon}{300}E_j+\frac{2\bar{c}_6}{\sigma^N}\omega(r_j)\lambda+\frac{2\bar{c}_7}{\sigma^N}\left[\frac{|\mu|(Q_{j-1})}{r_{j-1}^{N-1}}\right].\end{aligned} \tag{4.10}$$

Indeed, as in any case we have $E_j\leq 2\lambda$ and (4.5)-(4.6) hold, it follows $E_{j+1}\leq\varepsilon\lambda/50$ as a direct consequence of (4.10) when (4.9) does not hold. Notice that we have used that $(n+2)/[(p-1)n+p]\leq 1$, which holds since we are considering the case $p\geq 2$. Therefore to prove (4.8) we are reduced to check the validity of (4.10).

Step 2: Proof of (4.10). We note that the setting of the shrinking cylinders $\{Q_j\}$ and related comparison maps v_j,w_j defined in (4.7) is completely similar to the one adopted in the proof of Theorem 1.1, only the choice of the constants differs. To prove (4.10) is in turn sufficient to prove the following group of inequalities for the choices made in (4.3):

$$\sup_{\frac12 Q_j}|Dw_j|+s\leq A\lambda \tag{4.11}$$

$$\frac{\lambda}{A}\leq|Dw_{j-1}|\leq|Dw_{j-1}|+s\leq A\lambda \qquad \text{in } Q_j \tag{4.12}$$

$$\frac{\lambda}{B} \le \sup_{Q_{j+1}} |Dv_j| \le \sup_{\frac{1}{4}Q_j} |Dv_j| + s \le A\lambda \,. \tag{4.13}$$

Indeed, taking these for granted, let us see how to conclude with the proof of (4.10). Inequalities in displays (4.11)-(4.13) are completely analogous to (3.28) and (3.37)-(3.38), respectively. We can therefore exactly argue as in Step 6 of the proof of Theorem 1.1: using (4.11)-(4.12) we get (3.39) by Lemma 2.6, while (4.13) allows to use Theorem 2.4 with $\bar{\varepsilon} = 10^{-5N}\varepsilon$ thereby yielding

$$\fint_{Q_{j+1}} |Dv_j - (Dv_j)_{Q_{j+1}}|\, dx\, dt \le \frac{\varepsilon}{10^{5N}} \fint_{\frac{1}{4}Q_j} |Dv_j - (Dv_j)_{\frac{1}{4}Q_j}|\, dx\, dt$$

which plays the role of (3.40) in this context. Proceeding as in Step 6 we finally arrive at the inequality in display (4.10), which is the analog of (3.41) in this context.

It remains to prove (4.11)-(4.13). Using Lemma 2.2 and (4.5), we have

$$\begin{aligned}
&\left(\fint_{Q_{j+1}} |Du - Dw_{j-1}|^{p-1}\, dx\, dt\right)^{1/(p-1)} \\
&\qquad \le \sigma^{-\frac{2N}{p-1}} \left(\fint_{Q_{j-1}} |Du - Dw_{j-1}|^{p-1}\, dx\, dt\right)^{1/(p-1)} \le \frac{\lambda\varepsilon}{10^4}\,.
\end{aligned} \tag{4.14}$$

By (4.14), keeping (4.1) in mind, and using triangle inequality we get

$$\left(\fint_{Q_{j-1}} |Dw_{j-1}|^{p-1}\, dx\, dt\right)^{1/(p-1)} \le 2\lambda$$

so that, applying Corollary 2.1 to w_{j-1} we finally come to

$$\sup_{Q_j} |Dw_{j-1}| + s \le \sup_{\frac{1}{2}Q_{j-1}} |Dw_{j-1}| + s \le A\lambda \tag{4.15}$$

that proves the right hand side inequality in (4.12). In a completely similar way we obtain also (4.11). We now prove the left hand side inequality in (4.12). Again using (4.14), we get

$$\begin{aligned}
\sup_{Q_j} |Dw_{j-1}| \quad \ge \quad & \left(\fint_{Q_{j+1}} |Du|^{p-1}\, dx\, dt\right)^{1/(p-1)} \\
& - \left(\fint_{Q_{j+1}} |Du - Dw_{j-1}|^{p-1}\, dx\, dt\right)^{1/(p-1)} \ge \frac{\lambda\varepsilon}{50} - \frac{\lambda\varepsilon}{10^4} \ge \frac{\lambda\varepsilon}{100}.
\end{aligned}$$

The lower bound in the last display provides the existence of a point $(\tilde{x}, \tilde{t}) \in Q_j$ such that $|Dw_{j-1}(\tilde{x}, \tilde{t})| \ge \lambda\varepsilon/200$ while (4.15) allows to apply Theorem 2.3 (to w_{j-1} in $\frac{1}{2}Q_{j-1}$) getting, thanks to the choice of δ in (4.3),

$$\operatorname*{osc}_{Q_j} Dw_{j-1} \le \frac{\lambda\varepsilon}{10^5}\,.$$

The last two inequalities finally give $|Dw_{j-1}| \ge \lambda\varepsilon/10^5$ in Q_j. Summarizing, recalling the choice of A in (4.3) and again (4.15), the proof (4.12) turns out to be complete. It remains to prove (4.13). Using (4.11), Lemma 2.5 and finally (4.6) we have

$$\left(\fint_{\frac{1}{2}Q_j} |Dw_j - Dv_j|^{p-1}\, dx\, dt\right)^{1/(p-1)} \le \bar{c}_5^{1/p} \left[\omega(r_j)\right]^{2/p} \lambda \le \frac{\sigma^{3N}\lambda\varepsilon}{10^6}\,.$$

Combining the above estimate with (2.19) and (4.5) gives, after a few standard manipulations,

$$\max\left\{\left(\fint_{Q_{j+1}}|Du-Dv_j|^{p-1}\,dx\,dt\right)^{\frac{1}{p-1}},\left(\fint_{\frac12 Q_j}|Du-Dv_j|^{p-1}\,dx\,dt\right)^{\frac{1}{p-1}}\right\}$$

$$\leq \frac{\bar c_1}{\sigma^N}\lambda\left[\frac{1}{\lambda}\frac{|\mu|(Q_j)}{r_j^{N-1}}\right]^{(n+2)/[(p-1)n+p]}+\frac{\sigma^{2N}\lambda\varepsilon}{10^6}\leq\frac{\sigma^{2N}\lambda\varepsilon}{10^5}\,. \tag{4.16}$$

Therefore, using (4.16), Minkowski's inequality and recalling the definition of λ in (4.1) we have

$$\left(\fint_{\frac12 Q_j}|Dv_j|^{p-1}\,dx\,dt\right)^{1/(p-1)}\leq\lambda+\frac{\sigma^{2N}\lambda\varepsilon}{10^5}\leq 2\lambda$$

so that the right hand side inequality in (4.13) follows by Theorem 2.2. As for the left hand side, notice that the first inequality in (4.10) and (4.16) imply

$$\sup_{Q_{j+1}}|Dv_j|\geq\left(\fint_{Q_{j+1}}|Du|^{p-1}\,dx\,dt\right)^{1/(p-1)}$$
$$-\left(\fint_{Q_{j+1}}|Du-Dv_j|^{p-1}\,dx\,dt\right)^{1/(p-1)}\geq\frac{\lambda\varepsilon}{50}-\frac{\sigma^{2N}\lambda\varepsilon}{10^5}\geq\frac{\varepsilon\lambda}{10^5}\equiv\frac{\lambda}{B}$$

so that (4.13) is completely proved.

Step 3: Interpolation of radii. With (4.8) at our disposal we can finally conclude the proof of Theorem 1.4 by letting $r_\varepsilon:=\sigma^2\bar R_0$. Indeed, consider $\varrho\leq\sigma^2\bar R_0$; this means there exists an integer $m\geq 2$ such that $\sigma^{m+1}\bar R_0<\varrho\leq\sigma^m\bar R_0$. Therefore we have $\varrho=\sigma^m r$ for some $r\in(\sigma\bar R_0,\bar R_0]$ and (4.2) follows from (4.8) with this particular choice of r.

4.2. **Proof of Theorem 1.3.** The proof is now based on a combination of the arguments of Theorem 1.4 with those which are more typical of the elliptic case; we report everything in full detail for the sake of completeness and readability, and also because a certain number of modifications is really needed. We shall therefore keep the notation introduced in Step 1 of the proof of Theorem 1.4. Essentially, we are going to use the same choices in (4.3)-(4.6) but using an additional smallness condition on the radii used; in this way we can use both the inequalities in the proof of Theorem 1.4 and the result of Theorem 1.4. We consider a cylinder $Q_0\Subset\Omega_T$ and prove that for every $\varepsilon>0$ there exists a radius $r_\varepsilon\leq d_{\rm par}(Q_0,\partial\Omega_T)/2$, depending only on $n,p,\nu,L,\omega(\cdot),\mu(\cdot),\varepsilon$, such that

$$|(Du)_{Q^\lambda_\varrho(x_0,t_0)}-(Du)_{Q^\lambda_\rho(x_0,t_0)}|\leq\lambda\varepsilon\qquad\text{holds for every }\varrho,\rho\in(0,r_\varepsilon] \tag{4.17}$$

whenever $(x_0,t_0)\in Q_0$. This proves that Du is the local uniform limit of continuous maps - defined via the averages - and hence it is continuous. The rest of the proof goes in two steps.

Step 1: Dyadic sequences and continuity. To begin with the proof of (4.17) we recall that R_1 is determined in Theorem 2.1. Moreover R_3 is determined in (4.5)-(4.6) with the constant $\bar c_1,\bar c_5,\bar c_6$ and $\bar c_7$ obtained Theorem 1.4 and corresponding to the choices made in (4.3) and (4.4). Next, we take yet another positive radius $R_4\leq d_{\rm par}(Q_0,\partial\Omega_T)/2$ such that

$$\sup_{(x,t)\in Q_0}\int_0^{4R_4}\frac{|\mu|(Q_\varrho(x,t))}{\lambda\varrho^{N-1}}\,\frac{d\varrho}{\varrho}+\int_0^{4R_4}\omega(\varrho)\,\frac{d\varrho}{\varrho}\leq\frac{\sigma^{4N}\varepsilon}{10^6\bar c_6\bar c_7} \tag{4.18}$$

and, recalling the definition in (2.16)

$$\sup_{0<\varrho\leq R_4}\ \sup_{(x,t)\in Q_0} E(Du, Q_\varrho^\lambda(x,t)) \leq \frac{\sigma^{4N}\lambda\varepsilon}{10^5}\,. \tag{4.19}$$

Observe that it is possible to make the choice in (4.19) thanks to Theorem 1.4 and in particular to (4.2), that ensures that R_4 can be chosen in a way that makes it depending only on $n, p, \nu, L, \omega(\cdot), \mu(\cdot), \varepsilon$. Finally, this time we set $\tilde{R}_0 := \min\{R_1, R_3, R_4\}/2$ and take everywhere $r \leq \tilde{R}_0$ so that $\tilde{R}_0$ ultimately depends again on $n, p, \nu, L, \omega(\cdot), \mu(\cdot), \varepsilon$ only. The sequence of shrinking cylinders $\{Q_j\}$ is now defined as

$$Q_j \equiv Q_{r_j}^\lambda(x_0, t_0)\,, \qquad r_j = \sigma^j \tilde{R}_0, \quad \text{for} \ \ j \geq 0\,,$$

while $\lambda \geq 1$ is still defined as in (4.1) and σ in (4.4). By (4.18), computations similar to those in (3.10) and (3.15) then give

$$\sum_{i=0}^{\infty} \frac{|\mu|(Q_i)}{\lambda r_i^{N-1}} + \sum_{i=0}^{\infty} \omega(r_i) \leq \frac{\sigma^{3N}\varepsilon}{10^6 \bar{c}_6 \bar{c}_7}\,. \tag{4.20}$$

In Step 2 we will prove that

$$|(Du)_{Q_h} - (Du)_{Q_k}| \leq \frac{\lambda\varepsilon}{12} \qquad \text{holds whenever } 2 \leq k \leq h\,. \tag{4.21}$$

Here we show how to use (4.21) to finish the proof and to verify (4.17) with the choice $r_\varepsilon := \sigma^2 \tilde{R}_0$. Indeed, let us fix $0 < \rho < \varrho \leq r_\varepsilon$. This means that there exist two integers, $2 \leq k \leq h$, such that $\sigma^{k+1}\tilde{R}_0 < \varrho \leq \sigma^k \tilde{R}_0$ and $\sigma^{h+1}\tilde{R}_0 < \rho \leq \sigma^h \tilde{R}_0$. Applying (4.19) we get

$$\begin{aligned}
|(Du)_{Q_\varrho^\lambda(x_0,t_0)} - (Du)_{Q_{k+1}}| &\leq \fint_{Q_{k+1}} |Du - (Du)_{Q_\varrho^\lambda(x_0,t_0)}|\,dx\,dt \\
&\leq \frac{|Q_\varrho^\lambda(x_0,t_0)|}{|Q_{k+1}|} \fint_{Q_\varrho^\lambda(x_0,t_0)} |Du - (Du)_{Q_\varrho^\lambda(x_0,t_0)}|\,dx\,dt \\
&\leq \sigma^{-N} E(Du, Q_\varrho^\lambda(x_0,t_0)) \leq \frac{\lambda\varepsilon}{10}\,,
\end{aligned}$$

and, similarly,

$$|(Du)_{Q_\rho^\lambda(x_0,t_0)} - (Du)_{Q_{h+1}}| \leq \frac{\lambda\varepsilon}{10}\,.$$

Using the inequalities in the last two displays together with (4.21) and triangle inequality establishes (4.17) and the proof is complete, modulo the content of the next and final step.

Step 2: Proof of (4.21). The preliminary observation to make is that, with the choices made here, (4.10) holds in this setting as this ultimately relies on (4.5)-(4.6), that indeed are in force by the choice of $\tilde{R}_0$. To continue, let us consider the set $\mathcal{L}$ defined by

$$\mathcal{L} := \left\{ j \in \mathbb{N} \ : \ \left(\fint_{Q_j} |Du|^{p-1}\,dx\,dt \right)^{1/(p-1)} < \frac{\lambda\varepsilon}{50} \right\}\,,$$

and, accordingly, we then define the sets

$$\mathcal{C}_i^m = \{ j \in \mathbb{N} : i \leq j \leq i+m,\ i \in \mathcal{L},\ j \notin \mathcal{L} \text{ if } j > i \}$$

for $m \in \mathbb{N}$ and, finally, the number $j_e := \min \mathcal{L}$. Note that it may happen that $j_e = \infty$; in this case $\mathcal{L}$ is empty and the first inequality in (4.10) holds for every $j \geq 1$. The idea is now to employ (4.10) on suitable sets $\mathcal{C}_i^m$ using the indexes i as a sort of exit time indexes; the difference is that they can be countably many now.

We can now prove (4.21), obviously assuming $k < h$. The first case we analyze is when $k < h \le j_e$; we use (4.10) and the definition of j_e to infer that the inequality

$$E_{j+1} \le \frac{1}{2}E_j + \frac{2\bar{c}_6}{\sigma^N}\omega(r_j)\lambda + \frac{2\bar{c}_7}{\sigma^N}\left[\frac{|\mu|(Q_{j-1})}{r_{j-1}^{N-1}}\right] \tag{4.22}$$

holds for every $j \in \{k-1, \dots, h-2\}$. Summing up the previous inequalities easily yields

$$\sum_{i=k}^{h-1} E_i \le E_{k-1} + \frac{4\bar{c}_6}{\sigma^N}\sum_{j=0}^{\infty}\omega(r_j)\,\lambda + \frac{4\bar{c}_7}{\sigma^N}\sum_{j=0}^{\infty}\frac{|\mu|(Q_j)}{r_j^{N-1}} \le \frac{\sigma^{2N}\lambda\varepsilon}{50}$$

where we have used (4.19)-(4.20), therefore (4.21) follows since

$$\begin{aligned}
|(Du)_{Q_h} - (Du)_{Q_k}| &\le \sum_{i=k}^{h-1}|(Du)_{Q_{i+1}} - (Du)_{Q_i}| \\
&\le \sum_{i=k}^{h-1}\fint_{Q_{i+1}}|Du - (Du)_{Q_i}|\,dx\,dt \\
&\le \sum_{i=k}^{h-1}\frac{|Q_i|}{|Q_{i+1}|}\fint_{Q_i}|Du - (Du)_{Q_i}|\,dx\,dt \\
&= \sigma^{-N}\sum_{i=k}^{h-1}E_i \le \frac{\lambda\varepsilon}{50}\,. \qquad (4.23)
\end{aligned}$$

The second case we consider is when $j_e \le k < h$, where we prove (4.21) through the inequalities

$$|(Du)_{Q_h}| \le \frac{\lambda\varepsilon}{25} \qquad \text{and} \qquad |(Du)_{Q_k}| \le \frac{\lambda\varepsilon}{25}\,. \tag{4.24}$$

In (4.24), we prove the former, the argument for the latter being the same when $k > j_e$, otherwise $|(Du)_{Q_k}| \le \lambda\varepsilon/25$ is trivial if $k = j_e \in \mathcal{L}$. If $h \in \mathcal{L}$, the first inequality in (4.24) follows immediately from the definition of $\mathcal{L}$. On the other hand, if $h \notin \mathcal{L}$, then, as $h > j_e$, it is possible to consider a set $\mathcal{C}_{i_h}^{m_h}$ with $m_h > 0$, such that $h \in \mathcal{C}_{i_h}^{m_h}$; notice that $h > i_h$ as $h \notin \mathcal{L} \ni i_h$. Then (4.10) gives that (4.22) holds whenever $j \in \{i_h, \dots, i_h + m_h - 1\}$. Summing up and performing elementary manipulations gives

$$\sum_{i=i_h}^{i_h+m_h} E_i \le 2E_{i_h} + \frac{4\bar{c}_6}{\sigma^N}\sum_{j=0}^{\infty}\omega(r_j)\,\lambda + \frac{4\bar{c}_7}{\sigma^N}\sum_{j=0}^{\infty}\frac{|\mu|(Q_j)}{r_j^{N-1}} \le \frac{\sigma^{2N}\lambda\varepsilon}{50}$$

where again we have used (4.19)-(4.20). Therefore, as in (4.23), we have

$$|(Du)_{Q_h} - (Du)_{Q_{i_h}}| \le \sigma^{-N}\sum_{i=i_h}^{h-1}E_i \le \sigma^{-N}\sum_{i=i_h}^{i_h+m_h}E_i \le \frac{\lambda\varepsilon}{50}$$

and then, using that $|(Du)_{Q_{i_h}}| \le \lambda\varepsilon/50$ as $i_h \in \mathcal{L}$, we have

$$|(Du)_{Q_h}| \le |(Du)_{Q_{i_h}}| + |(Du)_{Q_h} - (Du)_{Q_{i_h}}| \le \frac{\lambda\varepsilon}{25}$$

that is (4.24). The last case to consider is when $k < j_e < h$, that can be actually treated by a combination of the first two. It suffices to prove that the inequalities in display (4.24) still hold. Indeed, the first inequality in (4.24) follows exactly as in the second case. As for the second estimate in (4.24), let us remark that, as $j_e \in \mathcal{L}$, we have that $|(Du)_{Q_{j_e}}| \le \lambda\varepsilon/50$. On the other hand, we can use the first

case $k < h \le j_e$ with $h = j_e$, thereby obtaining $|(Du)_{Q_{j_e}} - (Du)_{Q_k}| \le \lambda\varepsilon/50$ and therefore the second inequality in (4.24) follows via triangle inequality.

5. Proof of Theorem 1.5

The proof of Theorem 1.5 is a consequence of a few simple observations once Lemma 2.7 is at disposal and the sequence of comparison solutions w_j is introduced. Let's start by Theorem 1.1. Going back to Section 2.2, after having introduced the maps $\{w_j\}$ in Lemma 2.7 we can introduce the maps $\{v_j\}$ exactly as in (2.5). It is now easy to see that with this definition all the properties of the maps $\{w_j\}$ and $\{v_j\}$ described in Sections 2.2-2.4 and used in the proof of Theorem 1.1 hold, and especially Lemma 2.4-2.6. The only difference, which stems from the right hand sides in the inequalities in Lemma 2.7, is that the quantities $|\mu|(\lfloor Q_j \rfloor_{\mathrm{par}})$ appear instead of $|\mu|(Q_j)$; this is anyway irrelevant in the context of Theorem 1.1, in view of (3.13) and of the first equality in (3.15). The proof now follows exactly as in the finite energy case. As a consequence, all the corollaries of Theorem 1.1, starting by Theorem 1.2, hold for SOLA as well. Next, when passing to Theorems 1.3 and 1.4 the proofs remain completely the same upon using Lemma 2.7; note that this time the appearance of $|\mu|(\lfloor Q_j \rfloor_{\mathrm{par}})$ instead of $|\mu|(Q_j)$ gives no problem at all since all the proofs are based on the use of smallness conditions as (4.5) and (4.18).

Remark 5.1. Definition 1 of SOLA is quite natural, and is motivated by the standard way of approximating measures with bounded functions in the weak-* convergence, via parabolic smoothing, that is, using mollifiers that, although acting uniformly in space, act backward in time. The main difference with the standard elliptic case is that, instead of getting that

$$\limsup_h \, |\mu_h|(Q) \le |\mu|(\overline{Q}) \, ,$$

that is an inequality involving the full closure of Q, we get (1.23) so that the upper part of a cylinder does not play any role in the approximation. This is the advantage that allows to pass to the limit easily in the pointwse potential estimates. Moreover, this in accordance to the fact that, when dealing with evolutionary equations, the behavior at a certain instant of the solution only depends on what happened at past times, but not on what happens at that instant. Let us briefly recall the procedure. One fixes a family of smooth mollifiers (approximation of the identity) $\{\phi_h\}$ with $\phi_h := h^n \phi(x/h)$ with $\phi \in C_0^\infty(B_1)$, $\phi \ge 0$ and $\|\phi\|_{L^1} = 1$. Similarly, we consider another family of smooth mollifiers $\{\tilde{\phi}_h\}$, this time in one variable: $\tilde{\phi}_h := h\tilde{\phi}(x/h)$ with $\tilde{\phi} \in C_0^\infty((-1,1))$ and $\|\tilde{\phi}\|_{L^1} = 1$. We then define

$$\mu_h := \Big[\phi_h(\cdot)\tilde{\phi}_h(\cdot + 1/h)\Big] * \mu \in L^\infty \, ,$$

and notice that the mollification in the time variable is backward. In such a way we obtain a weakly* convergent sequence (in the sense of measures) and also (1.23). Such a sequence can be for instance used in [4] to derive the corresponding existence theorems. The point that we want to stress here is that, in order to give a more suitable definition of SOLA, it is very often not sufficient to take *any* approximation of the measure μ via weak* convergence (something that is for instance sufficient when deriving global estimates). Instead, a more careful way of approximating measures, tailored to both the geometry of the problem in question and the degree of fine properties of solutions one wants to derive, must be adopted.

Acknowledgements. The authors are supported by the ERC grant 207573 "Vectorial Problems" and by the Academy of Finland project "Potential estimates and applications for nonlinear parabolic partial differential equations". The authors also thank Paolo Baroni for remarks on a preliminary version of the paper.

References

[1] Acerbi E. & Mingione G.: Gradient estimates for a class of parabolic systems. *Duke Math. J.* 136 (2007), 285–320.

[2] Baroni P.: *New contributions to Nonlinear Calderón-Zygmund theory.* Ph. D. Thesis, Scuola Normale Superiore, Pisa 2012.

[3] Boccardo L. & Gallouët T.: Nonlinear elliptic and parabolic equations involving measure data. *J. Funct. Anal.* 87 (1989), 149–169.

[4] Boccardo L. & Dall'Aglio A. & Gallouët T. & Orsina L.: Nonlinear parabolic equations with measure data. *J. Funct. Anal.* 147 (1997), 237–258.

[5] Campanato S.: Equazioni paraboliche del secondo ordine e spazi $\mathfrak{L}^{2,\theta}(\Omega,\delta)$. *Ann. Mat. Pura Appl. (IV)* 73 (1966), 55-102.

[6] Cianchi A. & Maz'ya V.: Global Lipschitz regularity for a class of quasilinear equations. *Comm. PDE* 36 (2011),100–133.

[7] DiBenedetto E.: $C^{1+\alpha}$ local regularity of weak solutions of degenerate elliptic equations. *Nonlinear Anal.* 7 (1983), 827–850.

[8] DiBenedetto E.: On the local behaviour of solutions of degenerate parabolic equations with measurable coefficients. *Ann. Scu. Norm. Sup. Pisa Cl. Sci. (IV)* 13 (1986), 487–535.

[9] DiBenedetto E.: *Degenerate parabolic equations.* Universitext. Springer-Verlag, New York, 1993.

[10] DiBenedetto E. & Friedman A.: Hölder estimates for nonlinear degenerate parabolic systems. *J. reine ang. Math. (Crelles J.)* 357 (1985), 1–22.

[11] Driver B. K. & Gross L. & Saloff-Coste L.: Holomorphic functions and subelliptic heat kernels over Lie groups. *J. Eur. Math. Soc.* 11 (2009), 941–978.

[12] Duzaar F. & Mingione G.: Gradient estimates via non-linear potentials. *Amer. J. Math.* 133 (2011), 1093–1149.

[13] Duzaar F. & Mingione G.: Gradient estimates via linear and nonlinear potentials. *J. Funct. Anal.* 259 (2010), 2961–2998.

[14] Havin M. & Mazya V. G.: A nonlinear potential theory. *Russ. Math. Surveys* 27 (1972), 71–148.

[15] Hedberg L.I. & Wolff T.: Thin sets in nonlinear potential theory. *Ann. Inst. Fourier (Grenoble)* 33 (1983), 161–187.

[16] Jin T. & Mazya V. & Van Schaftingen J.: Pathological solutions to elliptic problems in divergence form with continuous coefficients. *C. R. Acad. Sci. Paris Ser. I* 347 (2009), 773-778.

[17] Kilpeläinen T. & Malý J.: Degenerate elliptic equations with measure data and nonlinear potentials. *Ann Scu. Norm. Sup. Pisa Cl. Sci. (V)* 19 (1992), 591–613.

[18] Kilpeläinen T. & Malý J.: The Wiener test and potential estimates for quasilinear elliptic equations. *Acta Math.* 172 (1994), 137–161.

[19] Kinnunen J. & Lewis J. L.: Higher integrability for parabolic systems of p-Laplacian type. *Duke Math. J.* 102 (2000), 253–271.

[20] Kinnunen J. & Lewis J. L.: Very weak solutions of parabolic systems of p-Laplacian type. *Ark. Mat.* 40 (2002), 105–132.

[21] Kinnunen J. & Lukkari T. & Parviainen M.: An existence result for superparabolic functions. *J. Funct. Anal.* 258 (2010), 713–728.

[22] Kinnunen J. & Lukkari T. & Parviainen M.: Local approximation of superharmonic and superparabolic functions in nonlinear potential theory. *Preprint 2012.*

[23] Kuusi T. & Mingione G.: Potential estimates and gradient boundedness for nonlinear parabolic systems. *Rev. Mat. Iberoamericana* 28 (2012), 535–576.

[24] Kuusi T. & Mingione G.: Nonlinear potential estimates in parabolic problems. *Rend. Lincei - Mat. e Appl.* 22 (2011), 161–174.

[25] Kuusi T. & Mingione G.: The Wolff gradient bound for degenerate parabolic equations. *J. Europ. Math. Soc.*, to appear.

[26] Kuusi T. & Mingione G.: Gradient regularity for nonlinear parabolic equations. *Ann Scu. Norm. Sup. Pisa Cl. Sci. (V)*, to appear.

[27] Kuusi T. & Mingione G.: A surprising linear type estimate for nonlinear elliptic equations. *C. R. Acad. Sci. Paris Ser. I* 349 (2011), 889–892.

[28] Kuusi T. & Mingione G.: Linear potentials in nonlinear potential theory. *Arch. Ration. Mech. Anal.* 207 (2013) 215–246.

[29] Kuusi T. & Mingione G.: New perturbation methods for nonlinear parabolic problems. *J. Math. Pures Appl. (IX)* 98 (2012), 390–427.

[30] Kuusi T. & Parviainen M.: Existence for a degenerate Cauchy problem. *manuscripta math.* 128 (2009), 213–249.

[31] Giusti E.: *Direct methods in the calculus of variations.* World Scientific Publishing Co., Inc., River Edge, NJ, 2003.
[32] Lindqvist P.: On the definition and properties of p-superharmonic functions. *J. reine angew. Math. (Crelles J.)* 365 (1986), 67–79.
[33] Lindqvist P.: Notes on the p-Laplace equation. *Univ. Jyväskylä, Report* 102 (2006).
[34] Lindqvist P. & Manfredi J.J.: Note on a remarkable superposition for a nonlinear equation. *Proc. AMS* 136 (2008), 133–140.
[35] Manfredi J.J.: Regularity for minima of functionals with p-growth. *J. Diff. Equ.* 76 (1988), 203–212.
[36] Manfredi J.J.: Regularity of the gradient for a class of nonlinear possibly degenerate elliptic equations. *Ph.D. Thesis.* University of Washington, St. Louis.
[37] Mingione G.: The Calderón-Zygmund theory for elliptic problems with measure data. *Ann Scu. Norm. Sup. Pisa Cl. Sci. (V)* 6 (2007), 195–261.
[38] Mingione G.: Gradient potential estimates. *J. Europ. Math. Soc.* 13 (2011), 459–486.
[39] Phuc N.C. & Verbitsky I. E.: Quasilinear and Hessian equations of Lane-Emden type. *Ann. of Math. (II)* 168 (2008), 859–914.
[40] Phuc N.C. & Verbitsky I. E.: Singular quasilinear and Hessian equations and inequalities. *J. of Funct. Anal.* 256 (2009), 1875–1906.
[41] Saloff-Coste L.: The heat kernel and its estimates. *Probabilistic approach to geometry, Adv. Stud. Pure Math.* 57, Math. Soc. Japan, Tokyo (2010), 405–436.
[42] Stein E. M.: Editor's note: the differentiability of functions in $\mathbb{R}^n$. *Ann. of Math. (II)* 113 (1981), 383–385.
[43] Stein E. M.: *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals.* Princeton Math. Series, 43. Princeton University Press, Princeton, NJ, 1993.
[44] Stein E. M. & Weiss G.: *Introduction to Fourier analysis on Euclidean spaces.* Princeton Math. Ser., 32. Princeton Univ. Press, Princeton, N.J. 1971.
[45] Trudinger N.S. & Wang X.J.: On the weak continuity of elliptic operators and applications to potential theory. *Amer. J. Math.* 124 (2002), 369–410.
[46] Urbano J.M.: *The method of intrinsic scaling. A systematic approach to regularity for degenerate and singular PDEs.* Lecture Notes in Mathematics, 1930. Springer-Verlag, Berlin, 2008.
[47] Vázquez J.L.: *Smoothing and decay estimates for nonlinear diffusion equations. Equations of porous medium type.* Oxford Lecture Series in Math. Appl., 33. Oxford University Press, Oxford, 2006. xiv+234 pp.

Tuomo Kuusi, Aalto University Institute of Mathematics, P.O. Box 111000 FI-00076 Aalto, Finland
E-mail address: tuomo.kuusi@tkk.fi

Dipartimento di Matematica e Informatica, Università di Parma, Parco Area delle Scienze 53/a, Campus, 43124 Parma, Italy
E-mail address: giuseppe.mingione@unipr.it.